

P30530 homo sapien  
Q00993 mus musculu

## ALIGNMENTS

Accession	Length	Species	Accession	Length	Species
44	495.5	UFO_HUMAN	887	10.1	homo sapiens
45	493	UFO_MOUSE	888	10.1	mus musculus

Scoring table: BLOSUM62

Scoring table: BLOSUM62

Searched: 77977 seqs. 28268293 residues

Database : SwissProt 37:\*

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	
1	4895	99.8	913	1	EDD1_HUMAN	O08345	homo sapien
2	4584.5	93.5	910	1	EDD1_RAT	O03474	rattus norv
3	4568	93.2	911	1	EDD1_MOUSE	O03146	mus musculus
4	2415	49.3	855	1	TRK3_HUMAN	Q16832	homo sapien
5	2413	49.2	854	1	TRK3_MOUSE	O62371	mus musculus
6	661	13.5	818	1	TRKB_CHICK	Q19187	gallus gall
7	659.5	13.5	821	1	TRKB_MOUSE	P15209	mus musculus
8	659.5	13.5	821	1	TRKB_RAT	O36604	rattus norv
9	658	13.4	822	1	TRKB_HUMAN	O16620	homo sapien
10	654.5	13.3	796	1	TRKA_HUMAN	O46529	homo sapien
11	653.5	13.3	779	1	TRKA_RAT	P35739	rattus norv
12	644.5	13.1	825	1	TRKC_CHICK	Q91009	gallus gall
13	642.5	13.1	825	1	TRKC_PIG	P24786	sus scrofa
14	633.5	12.9	839	1	TRKC_HUMAN	Q16288	homo sapien
15	614	12.5	852	1	TRKC_CHICK	Q10144	gallus gall
16	611.5	12.5	864	1	TRKC_RAT	Q03351	rattus norv
17	601.5	12.3	605	1	RTK2_GEOCY	P42159	geodia cyrdo
18	544.5	11.1	402	1	KROS_AVISU	P00529	avian sarco
19	544.5	11.1	507	1	KROS_CHICK	O08941	gallus gall
20	539.5	11.0	1363	1	ILPR_BRALA	O02466	branchiosto
21	535	10.9	2347	1	KROS_HUMAN	P08922	homo sapien
22	534.5	10.9	1382	1	INSR_HUMAN	P06213	homo sapien
23	530.5	10.8	1372	1	INSR_MOUSE	P15208	mus musculus
24	528	10.8	1520	1	ABL_DROME	P00522	drosophila
25	527.5	10.8	1383	1	INSR_RAT	P15127	rattus norv
26	525	10.7	1300	1	INSR_CAVPO	P14617	cavia porce
27	523	10.7	746	1	ABL_MLVB	P00521	abelson mur
28	523	10.7	1123	1	ABL_MOUSE	P00520	mus musculus
29	521.5	10.6	1379	1	MET_MOUSE	P16056	mus musculus
30	521	10.6	1268	1	IRR_HUMAN	P14616	homo sapien
31	520	10.6	1130	1	ABL1_HUMAN	P00519	homo sapien
32	520	10.6	2146	1	INSR_DROME	P09208	drosophila
33	518	10.6	1390	1	MET_HUMAN	P08581	homo sapien
34	517	10.5	2584	1	7LES_DROVI	P20806	drosophila
35	517	10.5	439	1	ABL_FSVHY	P10447	feline sarc
36	511	10.4	1182	1	ABL2_HUMAN	P42684	homo sapien
37	511	10.4	1367	1	IGIR_HUMAN	P08069	homo sapien
38	509	10.4	640	1	IGIR_BOVIN	Q05688	bos taurus
39	508.5	10.4	2554	1	7LES_DROME	P13368	drosophila
40	505.5	10.3	802	1	FGR4_HUMAN	P22455	homo sapien
41	505	10.3	1114	1	RET_HUMAN	P07949	homo sapien
42	503	10.3	1370	1	IGIR_RAT	P24062	rattus norv
43	497	10.1	1607	1	MP1R_LYMS	Q25410	lymnaea sta

RX MEDLINE; 95151638.  
 RA LAVA S., BUTLER R., SHELLING A.N., HANBY A.M., POULSON R.,  
 RA GANESAN T.S.;  
 RT "Isolation and characterization of an epithelial-specific receptor  
 tyrosine kinase from an ovarian cancer cell line";  
 RL CELL GROWTH DIFFER. 5:1173-1183(1994).  
 CC -1- FUNCTION: MAY BE INVOLVED IN CELL-CELL INTERACTIONS AND  
 RECOGNITION.  
 CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE - ADP +  
 PROTEIN TYROSINE PHOSPHATE.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED AT LOW LEVELS IN MOST ADULT  
 TISSUES AND IS HIGHEST IN THE BRAIN AND LUNG. ABUNDANT IN  
 BREAST CARCINOMA CELL LINES.  
 CC -1- DOMAIN: THE GLY/PRO-RICH DOMAINS MAY BE REQUIRED FOR AN UNUSUAL  
 GEOMETRY OF INTERACTION WITH LIGAND OR SUBSTRATES.  
 CC -1- ALTERNATIVE PRODUCTS: TWO ISOFORMS ARE PRODUCED BY ALTERNATIVE  
 SPLICING. THE LONG FORM (CALLED CAK I AND SHOWN HERE) AND THE  
 SHORT FORM (CALLED CAK II) ONLY DIFFER BY THE ABSENCE OF A 37  
 RESIDUES SEGMENT.  
 CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC  
 DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 F5/8 TYPE C DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-  
 PROTEIN KINASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L11315; G403387;  
 DR EMBL; L20817; G306475;  
 DR EMBL; G1754536;  
 DR EMBL; X98208; E251231;  
 DR EMBL; X99023; E251231; JOINED.  
 DR EMBL; X99024; E251231; JOINED.  
 DR EMBL; X99025; E251231; JOINED.  
 DR EMBL; X99026; E251231; JOINED.  
 DR EMBL; X99027; E251231; JOINED.  
 DR EMBL; X99028; E251231; JOINED.  
 DR EMBL; X99029; E251231; JOINED.  
 DR EMBL; X99030; E251231; JOINED.  
 DR EMBL; X99031; E251231; JOINED.  
 DR EMBL; X99032; E251231; JOINED.  
 DR EMBL; X99033; E251231; JOINED.  
 DR EMBL; X99034; E251231; JOINED.  
 DR EMBL; L57508; G1160925;  
 DR EMBL; X74979; G400463;  
 DR EMBL; X29093; G732800;  
 DR MIM; 600452;  
 DR MIM; 600408;  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; FALSE\_NEG.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS00111; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00239; RECEPTOR\_TYR\_KIN\_II; 1.  
 DR PROSITE; PS01285; FA58C.1; 1.  
 DR PROSITE; PS01286; FA58C.2; 1.  
 DR PFAM; PF00069; pkinase; 1.  
 DR PFAM; PF00754; F5\_F8\_type\_C; 1.  
 DR HSP; P11362; IFC1.  
 DR TRANSFERASE; TYROSINE-PROTEIN KINASE; GLYCOPROTEIN; SIGNAL;  
 KW PHOSPHORYLATION; TRANSMEMBRANE; RECEPTOR; ATP-BINDING;  
 KW ALTERNATIVE SPLICING.  
 FT SIGNAL 18  
 FT CHAIN 19 913  
 FT DOMAIN 19 416  
 FT TRANSMEM 417 443  
 FT DOMAIN 444 913  
 FT DOMAIN 31 185  
 FT DOMAIN 31 185  
 F5/8 TYPE C (PHOSPHOLIPID-BINDING,

FT DOMAIN 377 415  
 FT DOMAIN 476 601  
 FT DOMAIN 610 905  
 FT NP\_BIND 616 624  
 FT BINDING 655 655  
 FT ACT\_SITE 766 766  
 FT DISULFID 31 185  
 FT MOD\_RES 513 513  
 FT MOD\_RES 792 792  
 FT MOD\_RES 796 796  
 FT MOD\_RES 797 797  
 FT CARBOHYD 211 211  
 FT CARBOHYD 260 260  
 FT CARBOHYD 371 371  
 FT CARBOHYD 394 394  
 FT VARSPLIC 506 542  
 FT CONFLICT 94 94  
 FT CONFLICT 833 833  
 FT CONFLICT 847 867  
 SQ SEQUENCE 913 AA; 101127 MW; 16071364 CRC32;  
 Query Match 99.8%; Score 4895; DB 1; Length 913;  
 Best Local Similarity 99.9%; Pred. No. 1.9e-242;  
 Matches 912; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MGPEALSLILLLLVASGDADMKGHFDPKCRYALGMDRTIPDSDISASSWSDDTAAR 60  
 DB 1 MGPEALSLILLLLVASGDADMKGHFDPKCRYALGMDRTIPDSDISASSWSDDTAAR 60  
 QY 61 HSRLESSDGDGAWCPAGSVFPKEEYLQVLDLQRLHLVALVGTQGRHAGGLCKEFSRSL 120  
 DB 61 HSRLESSDGDGAWCPAGSVFPKEEYLQVLDLQRLHLVALVGTQGRHAGGLCKEFSRSL 120  
 QY 121 RYSDRGRRMGWKDRWGQEVISGNEDEPGVVLKDLGPPMVARLVRFPRADRVMSVCLRV 180  
 DB 121 RYSDRGRRMGWKDRWGQEVISGNEDEPGVVLKDLGPPMVARLVRFPRADRVMSVCLRV 180  
 QY 181 ELYGCLWRDGLLSYTPVGTQMYLSEAVYLDSTYDGHVGLQVGLQGLADGVVGLDD 240  
 DB 181 ELYGCLWRDGLLSYTPVGTQMYLSEAVYLDSTYDGHVGLQVGLQGLADGVVGLDD 240  
 QY 241 FRKSQELRWPCYDYVGNWSNHSFSSGYVEMEFEDRLRAFOAMQVHCNMHTLGAUPLGG 300  
 DB 241 FRKSQELRWPCYDYVGNWSNHSFSSGYVEMEFEDRLRAFOAMQVHCNMHTLGAUPLGG 300  
 QY 301 VECRRFRGPAMAWEGEPHMRHNLGGNLGDPRAVSVPLGGVRVARELQCRFLFAGPWLIFS 360  
 DB 301 VECRRFRGPAMAWEGEPHMRHNLGGNLGDPRAVSVPLGGVRVARELQCRFLFAGPWLIFS 360  
 QY 361 EISFTSDVYNNSSPALGGTFPPAPWPPGPPPTNFSSLELEPRGQOPVAKPEGSPTAILI 420  
 DB 361 EISFTSDVYNNSSPALGGTFPPAPWPPGPPPTNFSSLELEPRGQOPVAKPEGSPTAILI 420  
 QY 421 GCLVAIIILLIIIALMLWRLHWRLLSKAERVLEEBELTVHLSVPGDTILINNRPGPRE 480  
 DB 421 GCLVAIIILLIIIALMLWRLHWRLLSKAERVLEEBELTVHLSVPGDTILINNRPGPRE 480  
 QY 481 PPHYOEPRRGNPHSPACVPNGSALLSNPAYRLLATYARPPRPGPPTPAWAKPNT 540  
 DB 481 PPHYOEPRRGNPHSPACVPNGSALLSNPAYRLLATYARPPRPGPPTPAWAKPNT 540  
 QY 541 QAYSQDMEPEKPGAPLLPPPPQNSVPHYAEADIVTLOGVTGGNTYAVPALPPGAVGDP 600  
 DB 541 QAYSQDMEPEKPGAPLLPPPPQNSVPHYAEADIVTLOGVTGGNTYAVPALPPGAVGDP 600  
 QY 601 PRVDFPRSLRKEKLGEGQFGEVHLCVDSQDLVSLDFPLNVRKGHPLLVAVKILRPD 660  
 DB 601 PRVDFPRSLRKEKLGEGQFGEVHLCVDSQDLVSLDFPLNVRKGHPLLVAVKILRPD 660  
 QY 661 ATKNAERDNLKVKIMSRKLDPNRIILLGVCVQDDPDLCLMTIDYMGENDLQNLQSAHQLED 720

```

Db 661 ATNARNDFLKEVIMSRLLKPNILRLGVCVODDPLCMITDMENGDLNQFLSAHQED 720
QY 721 KAAEGAPGDCGAAAGGTTISYPMMLHVAQAQIAGSMRYLATLNFVHRDLATRNCLVGENFTI 780
Db 721 KAAEGAPGDCGAAAGGTTISYPMMLHVAQAQIAGSMRYLATLNFVHRDLATRNCLVGENFTI 780
QY 781 KIADFGMSRNLVAGDYRVOGRAVLPIRMWAEICILMGKFTTASDVWAFGVTLWEVLMC 840
Db 781 KIADFGMSRNLVAGDYRVOGRAVLPIRMWAEICILMGKFTTASDVWAFGVTLWEVLMC 840
QY 841 RAOPFGQLTDEQVIEVNAEGERDQGRQVLSRPPACQGLYELMLRCWSRESEQRPPFSQ 900
Db 841 RAOPFGQLTDEQVIEVNAEGERDQGRQVLSRPPACQGLYELMLRCWSRESEQRPPFSQ 900
QY 901 LHRFLAEDALNTV 913
Db 901 LHRFLAEDALNTV 913

RESULT 2
EDDI_RAT
ID EDDI_RAT STANDARD; PRT; 910 AA.
AC Q63474;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE EPITHELIAL DISCOIDIN DOMAIN RECEPTOR 1 PRECURSOR (EC 2.7.1.112)
DE (TYROSINE-PROTEIN KINASE CAK) (CELL ADHESION KINASE) (TYROSINE KINASE
DE DDK) (DISCOIDIN RECEPTOR TYROSINE KINASE) (PROTEIN-TYROSINE KINASE
DE PTK-3)
GN EDDRI OR PTK3.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE-BRAIN;
RX MEDLINE; 94173920.
RT SANCHEZ M.B., TAPLEY P., SAINI S.S., HE B., PULIDO D., BARBACID M.;
RT Multiple tyrosine protein kinases in rat hippocampal neurons;
RT Isolation of ptk-3, a receptor expressed in proliferative zones of
RT the developing brain;
RL PROC. NATL. ACAD. SCI. U.S.A. 91:1819-1823(1994).
CC !- FUNCTION: MAY BE INVOLVED IN CELL-CELL INTERACTIONS AND
CC RECOGNITION (BY SIMILARITY).
CC !- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC PROTEIN TYROSINE PHOSPHATE.
CC !- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC !- TISSUE SPECIFICITY: VARIOUS EMBRYONIC AND ADULT TISSUES; ALSO
CC PROLIFERATIVE ZONES OF THE DEVELOPING BRAIN; HIPPOCAMPAL NEURONS.
CC !- SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF TYR-PROTEIN
CC KINASES.
CC !- SIMILARITY: CONTAINS 1 F5/8 TYPE C DOMAIN.
CC !- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L26525; G432481;
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
DR PROSITE; PS01285; FA58C.1; 1.
DR PROSITE; PS01286; FA58C.2; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PFAM; PF00069; pkinase; 1.
DR PFAM; PF00754; F5_F8_type_C; 1.

```

```

DR HSP; P11362; IFGI.
KW TRANSFERASE; TYROSINE-PROTEIN KINASE; GLYCOPROTEIN; SIGNAL;
KW PHOSPHORYLATION; RECEPTOR; ATP-BINDING.
FT SIGNAL 1 19
FT CHAIN 20 910
FT DOMAIN 20 413
FT TRANSMEM 414 440
FT DOMAIN 441 910
FT DOMAIN 32 186
FT DOMAIN 378 412
FT DOMAIN 473 598
FT DOMAIN 607 902
FT NP_BIND 613 621
FT BINDING 652 621
FT ACT_SITE 763 763
FT DISULFID 32 186
FT MOD_RES 510 510
FT MOD_RES 789 789
FT MOD_RES 793 793
FT MOD_RES 794 794
FT CARBOHYD 212 212
FT CARBOHYD 261 261
FT CARBOHYD 371 371
FT CARBOHYD 391 391
SQ SEQUENCE 910 AA; 101164 MW; D69094A9 CRC32;

Query Match 93.5%; Score 4584.5; DB 1; Length 910;
Best Local Similarity 93.5%; Pred. No. 1.2e-226;
Matches 855; Conservative 20; Mismatches 34; Indels 5; Gaps 2;

QY 1 MGPEALSS-LLLLLLVASGDADMKGHFDPKACRYALGMQDRTIPDSDISASSWSSTAA 59
Db 1 MGTCTLSLLLLLLLVITIGDADMKGHFDPKACRYALGMQDRTIPDSDISASSWSSTAA 60
QY 60 RHRSLSSDGDGAWCPAGSVFPKEEYLVQDLRLHLVALVGTQGRHAGGLGFEFSRSYR 119
Db 61 RHRSLSSDGDGAWCPAGSVFPKEEYLVQDLRLHLVALVGTQGRHAGGLGFEFSRSYR 120
QY 120 LRVSRDGRRWKMDKRGQEVISGNEDEPGEVVKDGLPPMVARLRFYPRADRVMSVCLR 179
Db 121 LRVSRDGRRWKMDKRGQEVISGNEDEPGEVVKDGLPPMVARLRFYPRADRVMSVCLR 180
QY 180 VELYGCLWRDGLLSYAPVGTMYLSAVYLVNDSTVDGHTVGLYGGQLADGVVGLD 239
Db 181 VELYGCLWRDGLLSYAPVGTMYLSAVYLVNDSTVDGHTVGLYGGQLADGVVGLD 240
QY 240 DFRKSQELRVMPGYDYVGVNSHSFSSGYVEMEFEDRLRAFAQAMQVHCNNMHTLGARLPG 299
Db 241 DFRKSQELRVMPGYDYVGVNSHSFSSGYVEMEFEDRLRAFAQAMQVHCNNMHTLGARLPG 300
QY 300 GVECFRFRGTAMAWEGEPHRLNGLGNDPRARAVSVPLGGVRVARELQCRFLFAGPWLLF 359
Db 301 GVECFRFRGTAMAWEGEPHRLNGLGNDPRARAVSVPLGGVRVARELQCRFLFAGPWLLF 360
QY 360 SETSFISDVVNGSPALGGFFPAPWMPGPPPTNFSSLEPRGQOPVAKPKSSSTAIL 419
Db 361 SETSFISDVVNGSPALGGFFPAPWMPGPPPTNFSSLEPRGQOPVAKPKSSSTAIL 416
QY 420 IGCULVAILLLLLIALLMLRHLWRLLSKAERRVLEELTVHLSVPGDITLINNRPGPR 479
Db 417 IGCULVAILLLLLIALLMLRHLWRLLSKAERRVLEELTVHLSVPGDITLINNRPGPR 476
QY 480 EPPPYQEPFRGRNPHSAPCVNGSALLSNPAYRLLATYARPPRGPPPTPAWAKPTN 539
Db 477 EPPPYQEPFRGRNPHSAPCVNGSALLSNPAYRLLATYARPPRGPPPTPAWAKPTN 536
QY 540 TQAYSGDYMEPEKPGAPLPPPPQNSVPHVAEADIVTLOGVTGNTYAVPALPPGAVGDS 599
Db 537 TQAYSGDYMEPEKPGAPLPPPPQNSVPHVAEADIVTLOGVTGNTYAVPALPPGAVGDS 596
QY 600 PPRVDFPRSLRPFKEKLGEGQFGEVHLCVDSQDLVSLDFPLNVRKHGPHLLVAVKILRP 659

```

Db 597 PPRVDFPRSRURFKELGEGQFGEVHUCEVEDPDQDLVTSDFPISVQKGPPLLVAVKILRP 656  
 Qy 660 DATNARNDFELKEVKIMSRKLDNPNIRLLGVCVQDDPCLCMITDMYMGDLNQFLSAHOLE 719  
 Db 657 DATNARNDFELKEVKIMSRKLDNPNIRLLGVCVQDDPCLCMITDMYMGDLNQFLSAHOLE 716  
 Qy 720 DAAEAGPAGGQAAGPTTISYPMLLHVAAGIAGSMRYLATLNFVHRDLATRNCLVGENET 779  
 Db 717 NKVTOGLPGDRESQDPTTISYPMLLHVAAGIAGSMRYLATLNFVHRDLATRNCLVGENET 776  
 Qy 780 IKIADFGRNRLYAGDYRVRQGRAVLPIRMWAMECILLMGKFTTASDVWAFGVTLWEVLM 839  
 Db 777 IKIADFGRNRLYAGDYRVRQGRAVLPIRMWAMECILLMGKFTTASDVWAFGVTLWEVLM 836  
 Qy 840 CRAOPFGQLTDEQVIENAGAEFFRDQGRQVYLSRPPACPGQLYELMLRCWSRESEQRPPFS 899  
 Db 837 CRSOPFQLTDEQVIENAGAEFFRDQGRQVYLSRPPACPGQLYELMLRCWSRESEQRPPFS 896  
 Qy 900 QLHRELAEDALNTV 913  
 Db 897 QLHRELAEDALNTV 910  
 RESULT 3  
 EDDI\_MOUSE STANDARD; PRT; 911 AA.  
 AC Q03146;  
 DT 01-OCT-1994 (REL. 30, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE).  
 DE EPITHELIAL DISCOIDIN DOMAIN RECEPTOR 1 PRECURSOR (EC 2.7.1.112)  
 DE (TYROSINE-PROTEIN KINASE CAK) (CELL ADHESION KINASE) (TYROSINE KINASE  
 DE DDR) (DISCOIDIN RECEPTOR TYROSINE KINASE) (PROTEIN-TYROSINE KINASE  
 DE MPK-6).  
 GN EDDR1 OR CAK OR MPK6.  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
 [1]  
 RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
 RC STRAIN-C57BL;  
 RA MEDLINE; 96204002.  
 RA PEREZ J.L., JING S.Q., WONG T.W.;  
 RT "Identification of two isoforms of the Cak receptor kinase that are  
 RT coexpressed in breast tumor cell lines.";  
 RL ONCOGENE 12:1469-1477(1996).  
 RN [2]  
 RP SEQUENCE OF 766-822 FROM N.A.  
 RC STRAIN-C57BL; TISSUE=EMBRYONIC BRAIN;  
 RX MEDLINE; 93096484.  
 RA GILARDI-HEBENSTREIT P., NIETO M.A., FRAIN M., MATTEI M.-G.,  
 RA CHESTIER A., WILKINSON D.G., CHARNAY P.;  
 RT "An Eph-related receptor protein tyrosine kinase gene segmentally  
 RT expressed in the developing mouse hindbrain.";  
 RL ONCOGENE 7:2499-2506(1992).  
 CC - FUNCTION MAY BE INVOLVED IN CELL-CELL INTERACTIONS AND  
 CC RECOGNITION (BY SIMILARITY).  
 CC - CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE - ADP +  
 CC PROTEIN TYROSINE PHOSPHATE.  
 CC - SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC - TISSUE SPECIFICITY: THE PREDOMINANT ISOFORM CAK I, IS EXPRESSED IN  
 CC DEVELOPING EMBRYO AND ADULT BRAIN; CAK II IS EXPRESSED IN VARIOUS  
 CC EPITHELIAL CELLS.  
 CC - ALTERNATIVE PRODUCTS: TWO ISOFORMS ARE PRODUCED BY ALTERNATIVE  
 CC SPLICING. THE LONG FORM (CALLED CAK I AND SHOWN HERE) AND THE  
 CC SHORT FORM (CALLED CAK II) ONLY DIFFER BY THE ABSENCE OF A 37  
 CC RESIDUES SEGMENT.  
 CC - SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF TYR-PROTEIN  
 CC KINASES.  
 CC - SIMILARITY: CONTAINS 1 F5/8 TYPE C DOMAIN.  
 CC - SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-  
 CC PROTEIN KINASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 DR EMBL; L57509; G1161063; -  
 DR EMBL; X57240; G53198; -  
 DR PIR; S30502; S30502;  
 DR MGI; MGI:99216; CAK.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; FALSE\_NEG.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE; PS00239; RECEPTOR\_TYR\_KIN\_II; 1.  
 DR PROSITE; PS01285; FA58C\_1; 1.  
 DR PROSITE; PS01286; FA58C\_2; 1.  
 DR PFAM; PF00069; pkinase; 1.  
 DR PFAM; PF00754; F5\_F8\_type\_C; 1.  
 DR HSP; P11362; 1FGI.  
 KW TRANSFERASE; TYROSINE-PROTEIN KINASE; GLYCOPROTEIN; SIGNAL;  
 KW PHOSPHORYLATION; TRANSMEMBRANE; RECEPTOR; ATP-BINDING;  
 KW ALTERNATIVE SPLICING.  
 FT SIGNAL 1 19  
 FT CHAIN 20 911  
 FT DOMAIN 20 414  
 FT DOMAIN 415 441  
 FT DOMAIN 442 911  
 FT DOMAIN 32 186  
 FT DOMAIN 379 413  
 FT DOMAIN 474 599  
 FT DOMAIN 608 903  
 FT NP\_BIND 614 622  
 FT BINDING 653 653  
 FT ACT\_SITE 764 764  
 FT DISULFID 32 186  
 FT MOD\_RES 511 511  
 FT MOD\_RES 790 790  
 FT MOD\_RES 794 794  
 FT MOD\_RES 795 795  
 FT CARBOHYD 213 213  
 FT CARBOHYD 262 262  
 FT CARBOHYD 372 372  
 FT CARBOHYD 392 392  
 FT VARSPIC 503 539  
 FT SEQUENCE 911 AA; 101160 MW; 3DCBB321 CRC32;  
 SQ  
 Query Match 93.2%; Score 4568; DB 1; Length 911;  
 Best Local Similarity 93.3%; Pred. No. 8.3e-226;  
 Matches 854; Conservative 20; Mismatches 35; Indels 6; Gaps 3;  
 Qy 1 MPEALSS-LLLLLVASGDADMKHFDPAKCRVALGMDRTIPDSDISASSSSDSTAA 59  
 Db 1 MGTGLTSLLLLLLVLTIGDADMKHFDPAKCRVALGMDRTIPDSDISASSSSDSTAA 60  
 Qy 60 RHRSLSSDGDGAWCPAGVPKPEEYLQVDLRLHLVALVGTQGRHAGGLGKFFRSYR 119  
 Db 61 RHRSLSSDGDGAWCPAGVPKPEEYLQVDLRLHLVALVGTQGRHAGGLGKFFRSYR 120  
 Qy 120 LRYSDRGWRMGKDRWGQEVISGNEDEGVLKDLGPPMVARLVRFYPRADRVMSVCLR 179  
 Db 121 LRYSDRGWRMGKDRWGQEVISGNEDEGVLKDLGPPMVARLVRFYPRADRVMSVCLR 180  
 Qy 180 VELYGLNDRDGLLSYAPVQGMYSLEA-VYLNDSYDGHVVGGLQGLQLADGVVGL 238  
 Db 181 VELYGLNDRDGLLSYAPVQGMYSLEA-VYLNDSYDGHVVGGLQGLQLADGVVGL 240  
 Qy 239 DDFRKSQELRVWPGDYVYGVWSNHSFSSGVVMEFEFDRLRAFOAMQVHCNNHMTLGRLP 298  
 Db 239 DDFRKSQELRVWPGDYVYGVWSNHSFSSGVVMEFEFDRLRAFOAMQVHCNNHMTLGRLP 298





```
Db 241 DTQTHVHWPYDYGWRNESATNGIEMFEEDRINRTTKVHCNNMFAGVKIFK 300
Qy 300 GVECFRRGPAWAGEPMHNLGNLGDPRARAVSPLGGRVARFLQCRFLFAGPWLLF 359
Db 301 EVQCYF-RSEASEWEPNALSPLVLDVNPASRVTVPLHHRMASAIKCOVHFADTWMF 359
Qy 360 SEISFISD-VVNSSPALGGTFPPAPWPPGPPPTNFSSLEPRGQOPVAKPGSPAI 418
Db 360 SEITFQSDAAMYNSEAL-----PTSP-----MAPTTVDPMKLVDSNTRI 400
Qy 419 LGICVAILLLIALLMLRLHWRLLSKAERVLEEELTVHLSVPGDTILNRR-P 476
Db 401 LGICVAILFILLIIVILWRQWQKMLKASRRMLDDMTVSLSPSSFMFNRRSS 460
Qy 477 GPREP-----PPYQEP-PRGNPHSPACVPNGSALLSNPVRLLLATVARP 523
Db 461 SPSEQGSNTYDRIFFLPDPQEP-----SRLRKLPF-----495
Qy 524 PRGPGPPTPAWAKPTNTQAYSGDYWEPEKPGAPLLPPPPQNSVPHYAADIIVTLQVGTG 583
Db 495 -----APGEESGCGSVVQVPQSPG-----EGVPHYAADIIVNLQVGTG 535
Qy 584 NTYAVPALPPGAVGDPGRV-DEPRSLRREKELGEGQFGEVHLCEVDSQDLVSLDFPL 642
Db 536 NTYSPVAVIMDLGSKDVAVEEFPRKLLTFKELGEGQFGEVHLCEVGEKFKDKDFAL 595
Qy 643 NVRGHPLVAVKILRPDATKNARNDFLKEVKIMSLRKLDPNIIRLLGVCVQDDPLCMITD 702
Db 596 DVSANQPVAVKMLRADANKARNDFLKEIKIMSLRKLDPNIIRLLSVCTDPLCMITE 655
Qy 703 YWENGDLNQFLSAHQLEDKAEGAPGQQAAGQPTISYPMLLHVAAGIAGMRYLATLNF 762
Db 656 YWENGDLNQFLSRHE-----PPNSSSDVTVSYTNLKFMTQIAGSMKYLSSLNF 706
Qy 763 VHRDLATNCLVGNFTIKIADFGMSRLYAGDYRVQGRAVLPPIRWMAECILMKFTT 822
Db 707 VHRDLATNCLVGNFTIKIADFGMSRLYSGDIYRQGRAVLPPIRWMSIESILLGKFTT 766
Qy 823 ASDVWAFGVTLWELMLCAQPPGOLTDEQVIENAGFEFFRDQGRQVYLSRPPACPGOLYE 882
Db 767 ASDVWAFGVTLWETFTFCQEPYQSLSDEQVIENAGFEFFRDQGRQVYLSRPPACPDVYK 826
Qy 883 LMLKCSWRESQRPFPFSLHFLAE 907
Db 827 LMLSCWRDRDTKRNPSFOEIHLLLO 851

RESULT 5
ID TRK3_MOUSE STANDARD; PRT; 854 AA.
AC Q62371;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DE RECEPTOR PROTEIN-TYROSINE KINASE TKT PRECURSOR (EC 2.7.1.112)
DE (TYROSINE-PROTEIN KINASE TYRO 10) (NEUROTROPHIC TYROSINE KINASE,
DE RECEPTOR-RELATED 3).
CS NTRK3 OR TKT OR TYRO10.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94067796.
RA KARN T., HOLTRICH U., BRAEUNINGER A., BOEHME B., WOLF G.,
RA RUEBSAMEN-WAIGMANN H., STREBHARDT K.
RT "Structure, expression and chromosomal mapping of TKT from man and
RT mouse: a new subclass of receptor tyrosine kinases with a factor
RL viii-like domain."
RL ONCOGENE 8:3433-3440(1993).
RN [2]
RP SEQUENCE FROM N.A.
```

```
RC STRAIN-C57BL/6; TISSUE-BRAIN;
RX MEDLINE; 94151011.
RA LAI C., LEMKE G.E.;
RT "Structure and expression of the Tyro 10 receptor tyrosine kinase.";
RL ONCOGENE 9:877-883(1994).
CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE -
CC ADP + PROTEIN TYROSINE PHOSPHATE.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED; HIGH LEVELS IN SKELETAL
CC MUSCLE, HEART, CNS, AND KIDNEY; LESS IN OTHER TISSUES. THE MAJOR
CC 10 KD TRANSCRIPT IS EXPRESSED IN HIGH LEVELS IN HEART AND LUNG,
CC LESS IN BRAIN AND TESTIS.
CC -!- ALTERNATIVE PRODUCTS: DIFFERENT TRANSCRIPTS ARE DERIVED FROM ONE
CC GENE.
CC -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
CC DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 F5/8 TYPE C DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X76505; G435162; ALT_INIT.
CC MGD; MGI:97385; NTRK3.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
CC PROSITE; PS01285; FA58C.1; 1.
CC PROSITE; PS01286; FA58C.2; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC PFAM; PF00069; Pkinase; 1.
CC PFAM; PF00754; F5_P8_type_C; 1.
CC HSSP; P11362; 1FGI.
CC -----
CC TRANSFERASE; TYROSINE-PROTEIN KINASE; GLYCOPROTEIN; SIGNAL;
CC PHOSPHORYLATION; TRANSMEMBRANE; RECEPTOR; ATP-BINDING.
CC -----
CC SIGNAL 1 21
CC CHAIN 22 854
CC DOMAIN 22 399
CC TRANSMEM 400 421
CC DOMAIN 422 854
CC DOMAIN 30 185
CC DOMAIN 563 739
CC NP_BIND 569 577
CC BINDING 608 608
CC ACT_SITE 709 709
CC DISULFID 30 185
CC MOD_RES 739 739
CC CARBOHYD 121 121
CC CARBOHYD 213 213
CC CARBOHYD 261 261
CC CARBOHYD 280 280
CC CARBOHYD 372 372
CC SEQUENCE 854 AA; 96482 MW; B78B6551 CRC32;

Query Match 49.2%; Score 2413; DB 1; Length 854;
Best Local Similarity 52.3%; Pred. No. 3,4e-116;
Matches 481; Conservative 119; Mismatches 220; Indels 100; Gaps 15;

Qy 9 LLLLLLVASGDADMKGHEDPAKCYALGMQDRTTPDSDISASSWSSTAAHRSLSSD 68
:|||||: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 10 VLLLLLLILGSA--KAQVNPACRYPLGMSGGHIPDEDDITASSQWSESTAAYGRLDSE 67
69 GDGAWCPAGSVFPKE-EYELQVDLQRLHLVALVGTQGRHAGLGKFKFSRSLRYSDGR 127
||||||: | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 68 GDGAWCPPIPVQDDLEKFLDILTLHFTLVGTQGRHAGGIEFAPMKYNSRDGS 127
128 RWMGWKDRWGQEVISGNEDEPVGVLKDLGPPMVARLVRFYPRADRVMSVCLRVLYGLW 187
```

```

Db 128 RWSWRNRHGKQVLDGNSNPYDFLKDLEPPVAVRVLIPVTDHSMNVMCRVLYGCVW 187
QY 188 RGLLSYTPVGGTMYL--SEAVLYNDSTYDGHVGLGQYGLQGLADGVGLDDFRKSK 245
Db 188 LDGLVSYNAPAGQOQVLPGGSIYLDNSVDG-AGVYSMTGELGQLTDGVSLDDFTQTH 246
QY 246 ELRVWPGYDVGKSNHSFSGVYEMEFEDRLRAFOAMQOVHNNMHTLGLARLPGGVGCRF 305
Db 247 EYHWPGYDVGWRNESATNGEINFEEDRIETNMKVHCNNMFAGVKIFKEVQCYF 306
QY 306 RRGPMAMWEGEPNHRHNLGNLGDPRARAVSVPVGGVAREFLOCFLEAGPWLSEISFI 365
Db 307 -RSEASEWPTAVYFLVLDVNPVSARFVTVPLHHRMSAICQHFADTWNMFSEITFQ 365
QY 366 SD--VYNNSPALGGTTPAPMWPVPPPTNFSSLELEPRGQOPVAKPEGSTAILIGCL 423
Db 366 SDAAMYNNS--GALPTSP-----MAPTTDPMLKVDDSNTRILIGCL 405
QY 424 VAILLLLLIALLMLWRLHWRLLSKAERVLEELTVHLSVPGDTILNNR-----PGPR 479
Db 406 VALIFILLAIIVLMRWQWQKMLEKASRRMLDDEMTVSLPSESMFNNRRSSPSEQ 465
QY 480 EP-----PPYQEPNPRGNPHSAPCVPGNSGALLSNPAYRLLATVAREPPRGP 528
Db 466 ESNSTYDRIFPLRPDQEP-----SLRLKLPF-----495
QY 529 PPTPAWAKTNTQAVSGDYMEPEKPGAPLPPPPQNSVPHYAEADIVTLQGTGNTYAV 588
Db 495 -----APGEESGCGVVKPAQNGP-----EGVPHYAEADIVNLQGTGNTYCV 540
QY 589 PALPGAVGDGPRV--DFPRSLRKEKLGEGQFGEVHLCEVDSQDVLSDPLNVRKG 647
Db 541 PATMDLLSGKDVAVEEFPKRLKKEKLGEGQFGEVHLCEVEGMEKFKDKDFALDVAN 600
QY 648 HPLVAVKILRPDAVKANRDNFELKEVKIMSRKDPNIIRLLGVCVQDDPLCMITDYMENG 707
Db 601 QPVLVAVKMLRADANKNARNDFELKEIKMSRLKDNIRLLAVCTEDPLCMITEIMENG 660
QY 708 DLNQFLSAHQLEKAEAGAPGQAAQGTTSYPMLLHVAQAISGRMYLATLNFVHRDL 767
Db 661 DLNQFLSRHEPLSSGSSDA-----TVSYANKFMATQIASGRMYLSSLFVHRDL 710
QY 768 ATRNCLVGENETIKIADFGMSNLYAGDYRYVQGRAVPIRMAWECILMGKFTASDVW 827
Db 711 ATRNCLVGNKNTIKIADFGMSNLYSGDYRYIQGRAVPIRMSWESILLGKFTASDVW 770
QY 828 AFGVTLWEVLMCRAPQFQGLTDEQVIEAGFEFFRQGRQVYLSRPPACPGLYELMLRC 887
Db 771 AFGVTLWEVTFCCQEQPYSQLSDEQVIENTGEFFRQGRQIYLPQALCPDSVYKMLMSC 830
QY 888 WSRSESEPPFSQLHRFLAE 907
Db 831 WRRETKHRPSFOEIHLLIQ 850

RESULT 6
TRKB_CHICK
AC Q91987; Q91010; STANDARD; PRT; 818 AA.
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE BDNF / NT-3 GROWTH FACTORS RECEPTOR PRECURSOR (EC 2.7.1.112) (TRKB
DE TYROSINE KINASE) (TRKB-B).
OS GALLUS GALLUS (CHICKEN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;
OC NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE; 95047511.

```

```

RA VINH N., ERDMANN K., HEUMANN R.;
RT "Cloning and sequence analysis of a cDNA encoding a novel truncated
RT form of the chicken TrkB receptor.";
RL GENE 149:383-384(1994).
RP SEQUENCE FROM N.A.
RX MEDLINE; 94116452.
RA DECHANT G., BIFFO S., OKAZAWA H., KOLBECK R., POTTGIESSER J.,
RA BARDE Y.A.;
RT "Expression and binding characteristics of the BDNF receptor chick
RT trkB";
RL DEVELOPMENT 119:545-558(1993).
CC -1- FUNCTION: RECEPTOR FOR BRAIN-DERIVED NEUROTROPHIC FACTOR (BDNF),
CC NEUROTROPHIN-3 AND NEUROTROPHIN-4/5 BUT NOT NERVE GROWTH FACTOR
CC (NGF). INVOLVED IN THE DEVELOPMENT AND/OR MAINTENANCE OF THE
CC NERVOUS SYSTEM. THIS IS A TYROSINE-PROTEIN KINASE RECEPTOR. KNOWN
CC SUBSTRATES FOR THE TRK RECEPTORS ARE SHC, PI-3 KINASE, AND PLC-
CC GAMMA-1 (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +
CC PROTEIN TYROSINE PHOSPHATE.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- PTM: LIGAND-MEDIATED AUTO-PHOSPHORYLATION.
CC -1- ALTERNATIVE PRODUCTS: A TRUNCATED NON-CATALYTIC FORM WHICH
CC FUNCTION REMAINS STILL UNCLEAR IS PRODUCED BY ALTERNATIVE
CC SPLICING.
CC -1- SUBUNIT: EXISTS IN A DYNAMIC EQUILIBRIUM BETWEEN MONOMERIC (LOW
CC AFFINITY) AND DIMERIC (HIGH AFFINITY) STRUCTURES (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-
CC PROTEIN KINASES. HIGHEST SIMILARITY WITH OTHER TRK-TYPE RECEPTORS.
CC -1- SIMILARITY: THE REPEATED LEUCINE-RICH (LRR) SEGMENT IS FOUND IN
CC MANY PROTEINS. NUMBER IN THIS PROTEIN: 2.
CC -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS
CC TWO C2-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; X77251; G472934; -
DR EMBL; X77252; G472936; -
EMBL; X74109; G407799; -
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
DR PFAM; PF00047; ig; 1.
DR PFAM; PF00069; pkinase; 1.
DR PFAM; PF00560; LRR; 1.
DR HSP; P11362; IFGI.
KW TRANSFERASE; TYROSINE-PROTEIN KINASE; TRANSMEMBRANE; ATP-BINDING;
KW PHOSPHORYLATION; RECEPTOR; GLYCOPROTEIN; NEUROGENESIS; SIGNAL;
KW LEUCINE-REPEAT; REPEAT; ALTERNATIVE SPLICING; IMMUNOGLOBULIN FOLD.
FT SIGNAL 1 31
FT CHAIN 32 818
FT DOMAIN 32 426
FT TRANSMEM 427 450
FT DOMAIN 451 818
FT DOMAIN 71 116
FT REPEAT 71 92
FT REPEAT 95 116
FT DOMAIN 213 269
FT DOMAIN 300 364
FT DOMAIN 534 803
FT NP_BIND 540 548
FT BINDING 568 568
FT ACT_SITE 672 672
FT MOD_RES 512 512
FT MOD_RES 698 698
FT MOD_RES 702 702
FT BDNF / NT-3 GROWTH FACTORS RECEPTOR.
FT EXTRACELLULAR (POTENTIAL).
FT POTENTIAL.
FT CYTOPLASMIC (POTENTIAL).
FT LEUCINE-RICH REPEATS.
FT LRR 1.
FT LRR 2.
FT IG-LIKE C2-TYPE DOMAIN.
FT IG-LIKE C2-TYPE DOMAIN.
FT PROTEIN KINASE.
FT ATP (BY SIMILARITY).
FT ATP (BY SIMILARITY).
FT BY SIMILARITY.
FT PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT PHOSPHORYLATION (AUTO-) (BY SIMILARITY).

```

```
FT MOD_RES 703 703 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT FT 813 813 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT SITE 512 512 INTERACTION WITH SHC PROTEIN (BY
FT FT 512 512 SIMILARITY).
FT SITE 813 813 INTERACTION WITH PLC-GAMMA1 (BY
FT FT 813 813 SIMILARITY).
FT CARBOHYD 66 66 POTENTIAL.
FT FT 94 94 POTENTIAL.
FT CARBOHYD 120 120 POTENTIAL.
FT FT 120 120 POTENTIAL.
FT CARBOHYD 199 199 POTENTIAL.
FT FT 199 199 POTENTIAL.
FT CARBOHYD 204 204 POTENTIAL.
FT FT 204 204 POTENTIAL.
FT CARBOHYD 226 226 POTENTIAL.
FT FT 226 226 POTENTIAL.
FT CARBOHYD 253 253 POTENTIAL.
FT FT 253 253 POTENTIAL.
FT CARBOHYD 287 287 POTENTIAL.
FT FT 287 287 POTENTIAL.
FT CARBOHYD 324 324 POTENTIAL.
FT FT 324 324 POTENTIAL.
FT CARBOHYD 337 337 POTENTIAL.
FT FT 337 337 POTENTIAL.
FT CARBOHYD 408 408 POTENTIAL.
FT FT 408 408 POTENTIAL.
FT VARSPLIC 462 520 GPSVTSNDDSDASPLHISNGSNTPSSSEGGPDVAIIGMT
FT FT KIPDIEQVQFGLTNSQL -> VHGEVKGVLVDQIWLSQL
FT FT DCNDEQVMVTVNSDSTADNNRNLGFLVLFKIPLDG
FT FT (IN TRUNCATED FORM).
FT VARSPLIC 521 818 MISSING (IN TRUNCATED FORM).
FT FT 818 818 MISSING (IN TRUNCATED FORM).
SQ SEQUENCE 818 AA; 91736 MW; 63CE0454 CRC32;

Query Match 13.5%; Score 661; DB 1; Length 818;
Best Local Similarity 34.6%; Pred. No. 4.6e-27;
Matches 176; Conservative 71; Mismatches 149; Indels 112; Gaps 18;

Qy 446 LLSAERVLLEELTVH-LSPVGTILNRRPGPREPPVQEE---PRPGNPPH----- 496
Db 360 LLAKNEXGEDKRDVAFHMSVPGD-----GSGPIVDVDVYETTPNDLGDVTNNSQI 413
Qy 496 SAPCVPNG-----SALLSNPAYRLLLATYARPP-----RGPGPTPAWAKP 537
Db 414 TSPVSNKENEDSITVVVGVIALVCTGLVIMLLIKFGRHSKFGMKGP----- 464
Qy 538 TNOQAYSGDYMEPKPGAPLLPPPPQNSVPHYAEADITVLQGVGTNYAVPALPGAVG 597
Db 464 -SSVISND-----DDASAPL-----HHISNGSNTPSSSEGGPDVAIIGMTKIPVI 507
Qy 598 DGPPRVDPFRSLR-----FKEKLGEQFGEVHLCE---VDSPODLVSLDF 640
Db 508 ENQVFGTNSQLKPDPTFVQHKHNVKRELGEAGFVKVFLAECVNLCPQEQD----- 562
Qy 641 PLNVKRGHPLLVAVKILRPDATKNARNDLFKEVKIMSLKDPNIIRLLGVQVDDPLCMI 700
Db 562 -----KILVAVKTLK-DASDNARKDFHREALLNLQHEHIVKYGVCVEGDLIMV 612
Qy 701 TDYMENGLDNFLSAHQLEDKAAAGAPGQGAAG---PTISYPLLHVAQAASGMRYL 757
Db 613 FEYMKHGDNLKFLRAH-----GPDVAVLMAEGNRPAAELTQSMFLHQAQIAAGMYL 663
Qy 758 ATLNVFHRDLATNCLNGENFTIKIADGMSRNLVAGDYVRQGRAVLPTRWMAWECILM 817
Db 664 ASQHFVHRDLATNCLNGENLVLKIGDFGMGRDYSITDYIVGVGHTMLPRLNMPPEISIMY 723
Qy 818 KFTTASDVWAFGVTLLWVLMLCRAQPGQLTDEQVIENAGEFFRDQGRGVYLSRRPACP 877
Db 724 RKFTTESDVWSLGVVWLEIFTYGR-QPWYQLSNNEVIECI-----TOGR-VLQRPRTCP 775
Qy 878 QCLYELMLRCWSRESEQRPPFSQLHRFL 905
Db 776 KEVYDMLGCWQREPHMRNLNKEIHSLL 803

RESULT 7
TRKB_MOUSE
ID TRKB_MOUSE STANDARD; PRT; 821 AA.
AC P15209;
DT 01-APR-1990 (REL. 14, CREATED)
DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
```

```
DE DE BDNF / NT-3 GROWTH FACTORS RECEPTOR PRECURSOR (EC 2.7.1.112) (TRKB
DE TYROSINE KINASE) (GPI45-TRKB / GP95-TRKB) (TRK-B).
GN NTRK2 OR TRKB.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RN SEQUENCE FROM N.A. (GPI45).
RN TISSUE-BRAIN;
RN MEDLINE; 90059970.
RN KLEIN R., PARADA L.F., COULIER F., BARBACID M.;
RN "trkb, a novel tyrosine protein kinase receptor expressed during
RN mouse neural development";
RN EMBO J. 8:3701-3709(1989).
RN [2]
RN SEQUENCE FROM N.A. (GP95/TI).
RN TISSUE-BRAIN;
RN MEDLINE; 90263089.
RN KLEIN R., CONWAY D., PARADA L.F., BARBACID M.;
RN "The trkb tyrosine protein kinase gene codes for a second neurogenic
RN receptor that lacks the catalytic kinase domain.";
RN CELL 61:647-656(1990).
RN [3]
RN FUNCTION.
RN MEDLINE; 91249395.
RN SOPPET D., ESCANDON E., MARAGOS J., MIDDLEMAS D.S., REID S.W.,
RN BLAIR J., BURTON L.E., STANTON B.R., KAPLAN D.R., HUNTER T.,
RN NICOLICS K., PARADA L.F.;
RN "The neurotrophic factors brain-derived neurotrophic factor and
RN neurotrophin-3 are ligands for the trkb tyrosine kinase receptor.";
RN CELL 65:895-903(1991).
RN CC FUNCTION: RECEPTOR FOR BRAIN-DERIVED NEUROTROPHIC FACTOR (BDNF),
RN CC NEUROTROPHIN-3 AND NEUROTROPHIN-4/5 BUT NOT NERVE GROWTH FACTOR
RN CC (NGF). INVOLVED IN THE DEVELOPMENT AND/OR MAINTENANCE OF THE
RN CC NERVOUS SYSTEM. THIS IS A TYROSINE-PROTEIN KINASE RECEPTOR. KNOWN
RN CC SUBSTRATES FOR THE TRK RECEPTORS ARE SHC, PI-3 KINASE, AND PLC-
RN CC GAMMA-1.
RN CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE -> ADP +
RN CC PROTEIN TYROSINE PHOSPHATE.
RN CC -!- SUBUNIT: EXISTS IN A DYNAMIC EQUILIBRIUM BETWEEN MONOMERIC (LOW
RN CC AFFINITY) AND DIMERIC (HIGH AFFINITY) STRUCTURES.
RN CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
RN CC -!- PTM: LIGAND-MEDIATED AUTO-PHOSPHORYLATION.
RN CC -!- TISSUE SPECIFICITY: THE DIFFERENT FORMS ARE DIFFERENTIALLY
RN CC EXPRESSED IN VARIOUS CELL TYPES. THE T2 ISOFORM IS PRIMARILY
RN CC EXPRESSED IN NEURONS.
RN CC -!- ALTERNATIVE PRODUCTS: AT LEAST THREE FORMS ARE PRODUCED BY
RN CC ALTERNATIVE SPLICING; THESE ARE: GPI45-TRKB, T1 (GP95-TRKB) AND
RN CC T2.
RN CC -!- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-
RN CC PROTEIN KINASES. HIGHEST SIMILARITY WITH OTHER TRK-TYPE RECEPTORS.
RN CC -!- SIMILARITY: THE REPEATED LEUCINE-RICH (LRR) SEGMENT IS FOUND IN
RN CC MANY PROTEINS. NUMBER IN THIS PROTEIN: 2.
RN CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS
RN CC TWO C2-LIKE DOMAIN.
RN CC
RN CC This SWISS-PROT entry is copyright. It is produced through a collaboration
RN CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
RN CC the European Bioinformatics Institute. There are no restrictions on its
RN CC use by non-profit institutions as long as its content is in no way
RN CC modified and this statement is not removed. Usage by and for commercial
RN CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
RN CC or send an email to license@sib-sib.ch).
RN CC
RN CC EMBL; M33385; G202159; -
RN CC EMBL; X17647; G55506; -
RN CC PIR; S06943; S06943.
RN CC PIR; A35104; A35104.
RN CC MGD; MGI:97384; NTRK2.
RN CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
RN CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
RN CC PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
RN CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
```

Q H 506 . E S



CC -!- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-  
 CC PROTEIN KINASES. HIGHEST SIMILARITY WITH OTHER TRK-TYPE RECEPTORS.  
 CC -!- SIMILARITY: THE REPEATED LEUCINE-RICH (LRR) SEGMENT IS FOUND IN  
 CC MANY PROTEINS. NUMBER IN THIS PROTEIN: 2.  
 CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS  
 CC TWO C2-LIKE DOMAIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M52991; G207474; -;  
 CC EMBL; M52992; G207476; -;  
 CC EMBL; M52993; G207478; -;  
 CC PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 CC PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 CC PROSITE; PS00239; RECEPTOR\_TYR\_KIN\_II; 1.  
 CC PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 CC PROSITE; PF00047; ig; 1.  
 CC PFAM; PF00069; pkinase; 1.  
 CC PFAM; PF00560; LRR; 1.  
 CC HSP; P11362; IFGI.  
 CC TRANSFERASE; TYROSINE-PROTEIN KINASE; TRANSMEMBRANE; ATP-BINDING;  
 CC KW PHOSPHORYLATION; RECEPTOR; GLYCOPROTEIN; NEUROGENESIS; SIGNAL;  
 CC KW LEUCINE-REPEAT; REPEAT; IMMUNOGLOBULIN FOLD; ALTERNATIVE SPLICING.  
 CC SIGNAL 1 31  
 CC BY SIMILARITY.  
 CC FT CHAIN 32 821  
 CC FT DOMAIN 32 429  
 CC FT TRANSMEM 430 453  
 CC FT DOMAIN 454 821  
 CC FT DOMAIN 72 93  
 CC FT REPEAT 96 117  
 CC FT DOMAIN 214 270  
 CC FT DOMAIN 301 365  
 CC FT DOMAIN 537 806  
 CC FT NP\_BIND 543 551  
 CC FT BINDING 571 571  
 CC FT BINDING 571 571  
 CC FT MOD\_RES 515 515  
 CC FT MOD\_RES 701 701  
 CC FT MOD\_RES 705 705  
 CC FT MOD\_RES 706 706  
 CC FT MOD\_RES 816 816  
 CC FT SITE 515 515  
 CC  
 CC SITE 816 816  
 CC INTERACTION WITH PLC-GAMMA1 (BY  
 CC SIMILARITY).  
 CC FT CARBOHYD 57 816  
 CC FT CARBOHYD 95 95  
 CC FT CARBOHYD 121 121  
 CC FT CARBOHYD 178 178  
 CC FT CARBOHYD 205 205  
 CC FT CARBOHYD 241 241  
 CC FT CARBOHYD 254 254  
 CC FT CARBOHYD 280 280  
 CC FT CARBOHYD 325 325  
 CC FT CARBOHYD 338 338  
 CC FT CARBOHYD 411 411  
 CC FT VARSPPLIC 466 466  
 CC  
 CC FT VARSPPLIC 477 821  
 CC FT VARSPPLIC 466 474  
 CC FT VARSPPLIC 475 821  
 CC SEQUENCE 821 AA; 92186 MW; 884222D2 CRC32;

Query Match 13.5%; Score 659.5; DB 1; Length 821;  
 Best Local Similarity 30.4%; Pred. No. 5.5e-27;

Matches 201; Conservative 72; Mismatches 188; Indels 201; Gaps 22;

QY 373 SPALGGTTPPPAPWPPGPPPTNFSSLEPRGQOPVAKPEGSPAILIGLIVAILLLL 432  
 DB 217 SCVSGCDPLPLYWDVGNLVSKHNETSHTQSLRIINISDDSGKQICSV----- 268  
 QY 433 IFALMLRLHRRLLSKAERVLVEELTVHLSV---PGDTILINRRPGPREPP----- 483  
 DB 268 -----AENLVGEDQDSVNLTVHFAPTIFEL-----ESPTSDDHWC 302  
 QY 483 -PYQPRPRGNPPHAPCPVNGS-----ALLSNPA-----Y 513  
 DB 303 IPF---TVRGNPKPALQWYNGAILNESKYICTKIHTVNTHTYHGCLQDNPHTMNGDY 359  
 QY 514 RLL-----ATYARPPRGPGPTPAWAKPTNTQAYSGDYMEP--E 551  
 DB 360 TLMKNEYKDERQISAHFMGRPGVDYETNPYPVLYEDWTPTDI-----GDITNSNE 415  
 QY 552 KPGAPLLPPPPONSYPHYAEDIVTLQV-----581  
 DB 416 IPSTDVADQTNREHLSVYAVVYASVVGFCLLVMLLLKLARHSKFGMKGPASVISNDD 475  
 QY 581 -----TGGNTYAVPALPGAVDGP---PRVDEP-----RS 608  
 DB 476 SASPLHHISNGNTSPSSSEGGPDVILGKIPVIENQYFGITNSQLKPDFTVQHIKRH 535  
 QY 609 RLRFKEKLGEGQGEVHLCE---VDSQDVLSDPLPNVRKGHPLLVAVKILRPDATKNA 665  
 DB 536 NIVLRKELGEGAGFVLAECYNLCPEQ-----KILVAVKTLK-DASDNA 580  
 QY 666 RNDFLKEVKIMRLKDPNIIRLLGVCVDDPLCMITDYMENGDLNQFLSAHQLED-KAAE 724  
 DB 581 RKDFHREAELLNLQHEHIVKFYGVCEGDPILVMFYEMKKGDLNKLFLRAHGPDAVLMAE 640  
 QY 725 GAPGDAQAGGT-ISYPMLLHVAQIASGRYLAFLNFVHRDLATRNCLVGENETIKIA 783  
 DB 641 GNP-----PIELTQSQLHIAQIAGVYLASQHFVHRDLATRNCLVGENLLVIG 692  
 QY 784 DFGMSRLYAGDYRYVQGRAVLPIRMWAEICLMGKFTTASDVAFVTLWEVLMCLCRAQ 843  
 DB 693 DFGMSRDVYSTDYRYVGGHTMLPIRMWPPESIMYRKFTTESDWSLVGVVWEIFTYGR-Q 751  
 QY 844 PFQLTDEQVIENAGEFFRDQGRVYLSRPPACPOGLVELMLRCWSRESEQPPFSQHLR 903  
 DB 752 PMYQLSNNEVIECI-----TQGR--VLQPRTPCQEVYELMLGWCQREPHTRKNKNIHT 804  
 QY 904 FL 905  
 DB 805 LL 806

RESULT 9  
 ID TRKB\_HUMAN ..... STANDARD; PRT; 822 AA.  
 AC Q16620; Q16675;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE BDNF / NT-3 GROWTH FACTORS RECEPTOR PRECURSOR (EC 2.7.1.112) (TRKB  
 DE TYROSINE KINASE) (GP145-TRKB) (TRK-B).  
 GN NTRK2 OR TRKB  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 RN PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 RC [1]  
 RP SEQUENCE FROM N.A. (LONG FORM).  
 RC TISSUE-HIPPOCAMPUS;  
 RX MEDLINE; 9530922.  
 RA NAKAGAWARA A., LIU X.-G., IKEGAKI N., WHITE P.S., YAMASHIRO D.J.,  
 RA NYCUM L.M., BIEGEL J.A., BRODEUR G.M.;  
 RT "Cloning and chromosomal localization of the human TRK-B tyrosine  
 RT kinase receptor gene (NTRK2).";  
 RL GENOMICS 25:538-546(1995).

[2]  
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE (LONG AND TRUNCATED FORMS).  
TISSUE-BRAIN;  
MEDLINE; 95123473.  
RA SHELTON D.L., SUTHERLAND J., GRIPP J., CAMERATO T., ARMANINI M.P.,  
RX PHILLIPS H.S., CARROLL K., SPENCER S.D., LEVINSON A.D.;  
RT "Human trks: molecular cloning, tissue distribution, and expression  
of extracellular domain immunoadhesins.";  
RL J. NEUROSCI. 15:477-491(1995).  
[3]  
SEQUENCE FROM N.A. (TRUNCATED FORMS).  
TISSUE-HIPPOCAMPUS;  
MEDLINE; 95022162.  
RA ALLEN S.J., DAWBARN D., ECKFORD S.D., WILCOCK G.K., ASHCROFT M.,  
RX COLEBROOK S.M., FEENEY R., MACGOWAN S.H.;  
RT "Cloning of a non-catalytic form of human trkB and distribution of  
messenger RNA for trkB in human brain.";  
RL NEUROSCIENCE 60:825-834(1994).  
CC -!- FUNCTION: RECEPTOR FOR BRAIN-DERIVED NEUROTROPHIC FACTOR (BDNF),  
CC (NGF). INVOLVED IN THE DEVELOPMENT AND/OR MAINTENANCE OF THE  
CC NERVOUS SYSTEM. THIS IS A TYROSINE-PROTEIN KINASE RECEPTOR. KNOWN  
CC SUBSTRATES FOR THE TRK RECEPTORS ARE SHC, PI-3 KINASE, AND PLC-  
CC GAMMA-1.  
CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +  
CC PROTEIN TYROSINE PHOSPHATE.  
CC -!- SUBUNIT: EXISTS IN A DYNAMIC EQUILIBRIUM BETWEEN MONOMERIC (LOW  
CC AFFINITY) AND DIMERIC (HIGH AFFINITY) STRUCTURES (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC -!- PTM: LIGAND-MEDIATED AUTO-PHOSPHORYLATION.  
CC -!- ALTERNATIVE PRODUCTS: A TRUNCATED NON-CATALYTIC FORM WHICH  
CC FUNCTION REMAINS STILL UNCLEAR IS PRODUCED BY ALTERNATIVE  
CC SPLICING. THE RATIO OF FULL-LENGTH TO TRUNCATED FORM IS HIGHER IN  
CC FETAL THAN IN ADULT BRAIN.  
CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED, MAINLY IN THE NERVOUS  
CC TISSUE. IN THE CNS, EXPRESSION IS OBSERVED IN THE CEREBRAL CORTEX,  
CC HIPPOCAMPUS, THALAMUS, CHOROID PLEXUS, GRANULAR LAYER OF THE  
CC CEREBELLUM, BRAIN STEM, AND SPINAL CORD. IN THE PERIPHERAL NERVOUS  
CC SYSTEM, IT IS EXPRESSED IN MANY CRANIAL GANGLIA, THE OPHTHALMIC  
CC NERVE, THE VESTIBULAR SYSTEM, MULTIPLE FACIAL STRUCTURES, THE  
CC SUBMAXILLARY GLANDS, AND DORSAL ROOT GANGLIA.  
CC -!- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-  
CC PROTEIN KINASES. HIGHEST SIMILARITY WITH OTHER TRK-TYPE RECEPTORS.  
CC -!- SIMILARITY: THE REPEATED LEUCINE-RICH (LRR) SEGMENT IS FOUND IN  
CC MANY PROTEINS. NUMBER IN THIS PROTEIN: 2.  
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS  
CC TWO C2-LIKE DOMAIN.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: U12140; G530791; -;  
DR EMBL: S76473; G913718; -;  
DR EMBL: S76474; G913720; -;  
DR EMBL: X75958; G473008; -;  
DR MIM: 600456; -;  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE: PS00239; RECEPTOR\_TYR\_KIN\_II; 1.  
DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR PFAM: PF00047; ig; 1.  
DR PFAM: PF00069; pkinase; 1.  
DR PFAM: PF00560; LRR; 1.  
DR HSP: P11362; 1FG1.  
KW TRANSFERASE; TYROSINE-PROTEIN KINASE; TRANSMEMBRANE; ATP-BINDING;  
KW PHOSPHORYLATION; RECEPTOR; GLYCOPROTEIN; NEUROGENESIS; SIGNAL;  
KW LEUCINE-REPEAT; REPEAT; IMMUNOGLOBULIN FOLD; ALTERNATIVE SPLICING.  
FT SIGNAL 1 31

FT CHAIN	32	822	BDNF / NT-3 GROWTH FACTORS RECEPTOR.
FT DOMAIN	32	430	EXTRACELLULAR (POTENTIAL).
FT TRANSMEM	431	454	POTENTIAL.
FT DOMAIN	455	822	CYTOPLASMIC (POTENTIAL).
FT DOMAIN	72	117	LEUCINE-RICH REPEATS.
FT REPEAT	72	93	LRR 1.
FT REPEAT	96	117	LRR 2.
FT DOMAIN	214	270	IG-LIKE C2-TYPE DOMAIN.
FT DOMAIN	301	365	IG-LIKE C2-TYPE DOMAIN.
FT NP_BIND	538	807	PROTEIN KINASE.
FT BINDING	544	552	ATP (BY SIMILARITY).
FT ACT_SITE	572	572	BY SIMILARITY.
FT MOD_RES	676	676	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES	702	702	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES	706	706	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES	707	707	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES	817	817	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT SITE	516	516	INTERACTION WITH SHC PROTEIN (BY SIMILARITY).
FT SITE	817	817	INTERACTION WITH PLC-GAMMA-1 (BY SIMILARITY).
FT CARBOHYD	67	67	POTENTIAL.
FT CARBOHYD	95	95	POTENTIAL.
FT CARBOHYD	121	121	POTENTIAL.
FT CARBOHYD	178	178	POTENTIAL.
FT CARBOHYD	205	205	POTENTIAL.
FT CARBOHYD	241	241	POTENTIAL.
FT CARBOHYD	254	254	POTENTIAL.
FT CARBOHYD	280	280	POTENTIAL.
FT CARBOHYD	325	325	POTENTIAL.
FT CARBOHYD	338	338	POTENTIAL.
FT CARBOHYD	412	412	POTENTIAL.
FT VARSPIC	467	477	PASVISNDDDS -> FVLPHKIPLDG (IN TRUNCATED FORM).
FT VARSPIC	478	822	MISSING (IN TRUNCATED FORM).
SQ SEQUENCE	822 AA;	91998 MW;	9F96D971 CRC32;

## Query Match

13.4%; Score 658; DB 1; Length 822;

Best Local Similarity 47.4%; Pred. No. 6.5e-27;

Matches 144; Conservative 45; Mismatches 79; Indels 36; Gaps 9;

QY	607	RSRLREFKELGQGEVHLCE---	VDSQDLVSLDFPLNVRKGHPLLVAVKILRPDARK	663
DB	535	RHNVILKRELGEAGKVFLEACYNLCPOD-----	KILVAVKTLK-DASD	579
QY	664	NARDFELKVKIMSRKDPNIIRLLGVQVQDDPLCMITDYMENGDLNQFLSAHQLED-KA	722	
DB	580	NARKDFHREAEALLTNQHEHIVKFGVCGVEGDPIMVFEYMKHGLNKLFLRAHGPDAVLM	639	
QY	723	AEGAPGDGQAQOQPT-ISPMLLHVAAQTASGMRYLATLNFVHRDLATNCLVGENFTIK	781	
DB	640	AENP-----PTELTSQMLHIAQAAGMVIASQHFVHRDLATNCLVGENLLVK	691	
QY	782	IADFGMSRLNLYAGDYRVQGRAVLPPIRWNAWECILMGKFTTASDYWAFGVTLWEVLMCLR	841	
DB	692	IGDFGMSRDVSTDYRVGHTMPLPIRWMPPIESIMYRKFTTESDWSLGVLMWEIITFYTK	751	
QY	842	AQFGQLTDEQVIENAGFEFFRQGRVYLSRPPACQGLYELMLRCWSRESQRPPFSOL	901	
DB	752	-QPWYLSNNEVEICI-----TQGR-VLQRPRTCPQEVYELMLGCGWQREPHMRNKINGI	803	
QY	902	HREL 905		
DB	804	HTLL 807		

## RESULT 10

TRKA.HUMAN

ID TRKA.HUMAN

AC P04629; P08119;

DT 13-AUG-1987 (REL. 05, CREATED)

15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)  
15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
HIGH AFFINITY NERVE GROWTH FACTOR RECEPTOR PRECURSOR (EC 2.7.1.112)  
(TRK1 TRANSFORMING TYROSINE KINASE PROTEIN) (P140-TRKA) (TRK-A).  
GN NTRK1 OR TRK.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A. (TRKA-I ISOFORM).  
RC TISSUE-COLON;  
RX MEDLINE; 89181575.  
RA MARTIN-ZANCA D., OSKAM R., MITRA G., COPELAND T.D., BARBACID M.;  
RT "Molecular and biochemical characterization of the human trk proto-  
oncogene.";  
RL MOL. CELL. BIOL. 9:24-33(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE-BRAIN;  
RX MEDLINE; 95123473.  
RA SHELTON D.L., SUTHERLAND J., GRIPP J., CAMERATO T., ARMANINI M.P.,  
RA PHILLIPS H.S., CARROLL K., SPENCER S.D., LEVINSON A.D.;  
RT "Human trks: molecular cloning, tissue distribution, and expression  
of extracellular domain immunoadhesins.";  
RL J. NEUROSCI. 15:477-491(1995).  
RN [3]  
RP SEQUENCE OF 399-796 FROM N.A.  
RX MEDLINE; 86146854.  
RA MARTIN-ZANCA D., HUGHES S.H., BARBACID M.;  
RT "A human oncogene formed by the fusion of truncated tropomyosin and  
protein tyrosine kinase sequences.";  
RL NATURE 319:743-748(1986).  
RN [4]  
RP SEQUENCE OF 399-796 FROM N.A.  
RX MEDLINE; 88196074.  
RA KOZMA S.C., REDMOND S.M.S., SAURER S.M., GRONER B., HYNES N.E.;  
RT "Activation of the receptor kinase domain of the trk oncogene by  
recombination with two different cellular sequences.";  
RL EMBO J. 7:147-154(1988).  
RN [5]  
RP FUNCTION.  
RX MEDLINE; 91218846.  
RA HEMPSTEAD B.L., MARTIN-ZANCA D., KAPLAN D.R., PARADA L.F., CHAO M.V.;  
RT "High-affinity NGF binding requires coexpression of the trk proto-  
oncogene and the low-affinity NGF receptor.";  
RL NATURE 350:678-683(1991).  
RN [6]  
RP FUNCTION.  
RX MEDLINE; 91191557.  
RA KLEIN R., JING S., NANDURI V., O'ROURKE E., BARBACID M.;  
RT "The trk proto-oncogene encodes a receptor for nerve growth factor.";  
RL CELL 65:189-197(1991).  
RN [7]  
RP ALTERNATIVE SPLICING.  
RX MEDLINE; 93315496.  
RA BARKER P.A., LOMEN-HOERTH C., GENSCHE E.M., MEAKIN S.O., GLASS D.J.,  
RA SHOOTER E.M.;  
RT "Tissue-specific alternative splicing generates two isoforms of the  
trk receptor.";  
RL J. BIOL. CHEM. 268:15150-15157(1993).  
RN [8]  
RP VARIANT CIPA ARG-577.  
RX MEDLINE; 96331294.  
RA INDO Y., TSURUTA M., HAYASHIDA Y., KARIM M.A., OHTA K., KAWANO T.,  
RA MITSUBUCHI H., TONOKI H., AWAYA Y., MATSUDA I.;  
RT "Mutations in the TRKA/NGF receptor gene in patients with congenital  
insensitivity to pain with anhidrosis.";  
RL NAT. GENET. 13:485-488(1996).  
RN [9]  
RP MUTAGENESIS OF TYR-791.  
RX MEDLINE; 94179299.  
RA LOEB D.M., STEPHENS R.M., COPELAND T.D., KAPLAN D.R., GREENE L.A.;  
RT "A trk nerve growth factor (NGF) receptor point mutation affecting  
interaction with phospholipase C-gamma 1 abolishes NGF-promoted  
peripherin induction but not neurite outgrowth.";  
J. BIOL. CHEM. 269:8901-8910(1994).  
RN [10]  
RP MUTAGENESIS, AND PHOSPHORYLATION SITES.  
RX MEDLINE; 94206546.  
RA STEPHENS R.M., LOEB D.M., COPELAND T.D., PAMSON T., GREENE L.A.,  
RA KAPLAN D.R.;  
RT "Trk receptors use redundant signal transduction pathways involving  
SHC and PLC-gamma 1 to mediate NGF responses.";  
RL NEURON 12:691-705(1994).  
RN [11]  
RP STRUCTURE BY NMR OF 489-500.  
RX MEDLINE; 96097066.  
RA ZHOU M.M., RAVICHANDRAN K.S., OLEJNICZAK E.F., PETROS A.M.,  
RA MEADOWS R.P., SATTLER M., HARLAN J.E., WADE W.S., BURAKOFF S.J.,  
RA PESIK S.W.;  
RT "Structure and ligand recognition of the phosphotyrosine binding  
domain of Shc.";  
RL NATURE 378:584-592(1995).  
RN [12]  
RP FUNCTION: REQUIRED FOR HIGH-AFFINITY BINDING TO NERVE GROWTH  
FACTOR (NGF). NEUROTROPHIN-3 AND NEUROTROPHIN-4/5 BUT NOT BRAIN-  
DERIVED NEUROTROPHIC FACTOR (BDNF). KNOWN SUBSTRATES FOR THE TRK  
RECEPTORS ARE SHC, PI-3 KINASE, AND PLC-GAMMA-1. HAS A CRUCIAL  
ROLE IN THE DEVELOPMENT AND FUNCTION OF THE NOCICEPTIVE RECEPTION  
SYSTEM AS WELL AS ESTABLISHMENT OF THERMAL REGULATION VIA  
SWEATING. ACTIVATES ERK1 BY EITHER SHC- OR PLC-GAMMA-1-DEPENDENT  
SIGNALING PATHWAY.  
CC CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +  
PROTEIN TYROSINE PHOSPHATE.  
CC SUBUNIT: EXISTS IN A DYNAMIC EQUILIBRIUM BETWEEN MONOMERIC (LOW  
AFFINITY) AND DIMERIC (HIGH AFFINITY) STRUCTURES.  
CC CELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC PTM: LIGAND-MEDIATED AUTO-PHOSPHORYLATION.  
CC ALTERNATIVE PRODUCTS: TWO ISOFORMS, TRKA-I AND TRKA-II (SHOWN  
HERE), ARE PRODUCED BY ALTERNATIVE SPLICING. BOTH ISOFORMS HAVE  
SIMILAR BIOLOGICAL PROPERTIES.  
CC TISSUE SPECIFICITY: TRKA-II IS PRIMARILY EXPRESSED IN NEURONAL  
CELLS; TRKA-I IS FOUND IN NON-NEURONAL TISSUES.  
CC DISEASE: THE AMINO END OF SEVERAL DIFFERENT CELLULAR PROTEINS  
(SUCH AS NON-MUSCLE TROPOMYOSIN TM30NM) HAS BEEN FOUND FUSED WITH  
THE PROTEIN TYROSINE KINASE DOMAIN ENCODED BY NTRK1. THESE SOMATIC  
REARRANGEMENT CREATES AN ONCOGENIC PROTEIN.  
CC DISEASE: DEFECTS IN NTRK1 ARE A CAUSE OF A DISEASE CALLED CIPA.  
CC WHICH IS CHARACTERIZED BY A CONGENITAL INSENSITIVITY TO PAIN AND  
BY ANHIDROSIS (ABSENCE OF SWEATING).  
CC SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-  
PROTEIN KINASES. HIGHEST SIMILARITY WITH OTHER TRK-TYPE RECEPTORS.  
CC SIMILARITY: THE REPEATED LEUCINE-RICH (LRR) SEGMENT IS FOUND IN  
MANY PROTEINS. NUMBER IN THIS PROTEIN: 2.  
CC SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS  
TWO C2-LIKE DOMAIN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M23102; G339918;  
CC EMBL; X03541; G37403; ALT\_INIT.  
CC EMBL; X06704; G37400; ALT\_INIT.  
CC PIR; A30124; TVHUTT.  
CC PDB; 1SHC; 15-MAY-97.  
CC MIM; 191315;  
CC MIM; 164970;  
CC MIM; 256800;  
CC PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
CC PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
CC PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
CC PROSITE; PS00239; RECEPTOR\_TYR\_KIN\_II; 1.



DR PFAM; PF00069; pkinase; 1.  
 DR PFAM; PF00560; LRR; 2.  
 KW TRANSFERASE; TYROSINE-PROTEIN KINASE; ATP-BINDING; PHOSPHORYLATION;  
 KW RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL; REPEAT; LEUCINE-REPEAT;  
 KW PROTO-ONCOGENE; CHROMOSOMAL TRANSLOCATION; NEUROGENESIS; 3D-STRUCTURE;  
 KW IMMUNOGLOBULIN FOLD; DISEASE MUTATION; ALTERNATIVE SPLICING.  
 FT SIGNAL 1 32  
 FT CHAIN 33 796  
 FT  
 FT DOMAIN 33 423  
 FT TRANSMEM 424 439  
 FT DOMAIN 440 796  
 FT DOMAIN 72 117  
 FT REPEAT 72 93  
 FT REPEAT 96 117  
 FT REPEAT 117 269  
 FT DOMAIN 211 269  
 FT DOMAIN 299 365  
 FT DOMAIN 510 781  
 FT NP\_BIND 516 524  
 FT BINDING 544 544  
 FT ACT\_SITE 650 650  
 FT MOD\_RES 496 496  
 FT MOD\_RES 676 676  
 FT MOD\_RES 680 680  
 FT MOD\_RES 681 681  
 FT MOD\_RES 791 791  
 FT SITE 392 393  
 FT  
 FT SITE 496  
 FT SITE 791 791  
 FT MUTAGEN 496 496  
 FT MUTAGEN 544 544  
 FT MUTAGEN 791 791  
 FT  
 FT CARBOHYD 67 67  
 FT CARBOHYD 95 95  
 FT CARBOHYD 121 121  
 FT CARBOHYD 188 188  
 FT CARBOHYD 202 202  
 FT CARBOHYD 253 253  
 FT CARBOHYD 262 262  
 FT CARBOHYD 281 281  
 FT CARBOHYD 318 318  
 FT CARBOHYD 323 323  
 FT CARBOHYD 338 338  
 FT CARBOHYD 358 358  
 FT CARBOHYD 401 401  
 FT VARSPLIC 393 398  
 FT VARIANT 577 577  
 FT CONFLICT 300 300  
 Query Match 13.3%; Score 654.5; DB 1; Length 796;  
 Best Local Similarity 30.3%; Pred. No. 9.5e-27;  
 Matches 185; Conservative 73; Mismatches 151; Indels 201; Gaps 19;  
 QY 349 RFLFAGPMLFSEISFI-----SDVYNSPALGGTTP----- 382  
 DB 314 RFLFNGS--VLNFSFIETEFLEPAANETVRHGCLRLNQPTHVNGNYTLIAANPFQOAS 371  
 QY 382 -----PAPWMPGPPPTNFSSLEPRGQOPVAKPGESPTAILGCLVAIILL 431  
 DB 372 ASIMAAFMNPFENPEDIPVSPVDTSNGDPEVKDETFFGVSVAVLAVFACLF 431  
 QY 432 LIALMLRLHWRLLSKAER-----VLTEE-----LTVHLSVPGDTILNNRPG 478  
 DB 432 LSTLLL-----VLNKCGRNFGINRPAVLAPEDGLAMSLHMTLGSSL----- 477  
 QY 479 REPPYQEPGRGNPHSPACVPNGSAL--LLSNPAYRLLLIATYARPGPGPTPAWA 535  
 DB 477 -----SPTGEGSGLQGHIIENPOY----- 497  
 QY 536 KPTNTQAYSGDMEPEKFGAPLLPPPPQNSVPHYAEADIVTLQGVTCGNTYAVPALPPGA 595

DB 497 -----FS-----DACHVHIKRDIV----- 512  
 QY 596 VGDGPPRVDFPRSLRPFKEKLGEGQFGEVHLCEDVS---PQDLVSLDFPLNVRKHGHPLLV 652  
 DB 512 -----LKWELGEGAFCKVFLAECHNLLPEQD-----KMLV 541  
 QY 653 AVKILRPDANKARNDFLKEVKIMSLKDNINIRLLGVCVQDDPLCLMITYMENGDLNQF 712  
 DB 542 AVKALK-EASESARQDFORAEALLTMLQHQHIVRFECVCTEGRPLLMVFYMRHGLNRF 600  
 QY 713 LSAHQLEKAAEGAPGDGQAAGPTISYPMLLHVAQOIASSMRYLATLNEVHRLATRNC 772  
 DB 601 LRSRGPDAKLAGE---DVAPGP-ICLGOLLAVASOVAAGMVLGLHFEVHRLATRNC 656  
 QY 773 LVGENFTIKIADFGMSNLVAGDYRVQGRAVLPIRMWAEICILMGKFTTASDVAFGVT 832  
 DB 657 LVGQGLVVKIGDFGMSRDIYSTDYRVVGRGTMLPIRMWPPESILYRKFTTESDVWSEGVV 716  
 QY 833 LWEVLMCRAQPFQGLTDEQVIENAGEFFRDQGRQVYLSRPPACOGLYELMLRCSRES 892  
 DB 717 LWEFTYTK-OPWYQLSNTEAIDCI-----TOGRE--LERPRACPEVYAIMRCWQREP 768  
 QY 893 EORPPFSQLH 902  
 DB 769 QQRHSIKDVH 778  
 RESULT 11  
 TRKA\_RAT  
 ID TRKA\_RAT STANDARD; PRT; 799 AA.  
 AC P35739;  
 DT 01-JUN-1994 (REL. 29, CREATED)  
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE HIGH AFFINITY NERVE GROWTH FACTOR RECEPTOR PRECURSOR (EC 2.7.1.112)  
 DE (P140-TRKA) (SLOW NERVE GROWTH FACTOR RECEPTOR) (TRKA-A).  
 GN NTRK1 OR TRKA OR TRK.  
 OS RATTUS NORVEGICUS (RAT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
 RN [1]  
 RN SEQUENCE FROM N.A. (TRKA-II ISOFORM).  
 RX MEDLINE; 92196121.  
 RA MEAKIN S.O., SUTER U., DRINKWATER C.C., WELCHER A.A., SHOOTER E.M.;  
 RT "The rat trk protooncogene product exhibits properties characteristic  
 RT of the slow nerve growth factor receptor.";  
 RL PROC. NATL. ACAD. SCI. U.S.A. 89:2374-2378(1992).  
 RN [2]  
 RN ALTERNATIVE SPLICING.  
 RX MEDLINE; 93315496.  
 RA BARKER P.A., LOMEN-HOERTH C., GENSCH E.M., MEAKIN S.O., GLASS D.J.,  
 RA SHOOTER E.M.;  
 RT "Tissue-specific alternative splicing generates two isoforms of the  
 RT trka receptor.";  
 RL J. BIOL. CHEM. 268:15150-15157(1993).  
 CC -!- FUNCTION: REQUIRED FOR HIGH-AFFINITY BINDING TO NERVE GROWTH  
 CC FACTOR (NGF). NEUROTROPHIN-3 AND NEUROTROPHIN-4/5 BUT NOT BRAIN-  
 CC DERIVED NEUROTROPHIC FACTOR (BDNF). KNOWN SUBSTRATES FOR THE TRK  
 CC RECEPTORS ARE SHC, PI-3 KINASE, AND PLC-GAMMA-1. HAS A CRUCIAL  
 CC ROLE IN THE DEVELOPMENT AND FUNCTION OF THE NOCICEPTIVE RECEPTION  
 CC SYSTEM AS WELL AS ESTABLISHMENT OF THERMAL REGULATION VIA  
 CC SWATING. ACTIVATES ERK1 BY EITHER SHC- OR PLC-GAMMA-1-DEPENDENT  
 CC SIGNALING PATHWAY (BY SIMILARITY).  
 CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +  
 CC PROTEIN TYROSINE PHOSPHATE.  
 CC -!- SUBUNIT: EXISTS IN A DYNAMIC EQUILIBRIUM BETWEEN MONOMERIC (LOW  
 CC AFFINITY) AND DIMERIC (HIGH AFFINITY) STRUCTURES.  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -!- PTM: LIGAND-MEDIATED AUTO-PHOSPHORYLATION.  
 CC -!- ALTERNATIVE PRODUCTS: TWO ISOFORMS, TRKA-I AND TRKA-II (SHOWN  
 CC HERE). ARE PRODUCED BY ALTERNATIVE SPLICING. BOTH ISOFORMS HAVE  
 CC SIMILAR BIOLOGICAL PROPERTIES.



CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC -!- SUBUNIT: EXISTS IN A DYNAMIC EQUILIBRIUM BETWEEN MONOMERIC (LOW  
CC AFFINITY) AND DIMERIC (HIGH AFFINITY) STRUCTURES (BY SIMILARITY).  
CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN THE CONDENSING DORSAL ROOT  
CC GANGLIA AT EMBRYONAL DAY 3, AND IN THE PRIMARY SYMPATHETIC CHAIN  
CC GANGLIA AT EMBRYONAL DAY 4.  
CC -!- ALTERNATIVE PRODUCTS: TWO ISOFORMS ARE PRODUCED BY ALTERNATIVE  
CC SPLICING OF THE SAME GENE.  
CC -!- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-  
CC PROTEIN KINASES. HIGHEST SIMILARITY WITH OTHER TRK-TYPE RECEPTORS.  
CC -!- SIMILARITY: THE REPEATED LEUCINE-RICH (LRR) SEGMENT IS FOUND IN  
CC MANY PROTEINS. NUMBER IN THIS PROTEIN: 2.  
CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS  
CC TWO C2-LIKE DOMAIN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: X93581; E213280; -  
CC PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
CC PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
CC PROSITE: PS00111; PROTEIN\_KINASE\_DOM; 1.  
CC PROSITE: PS00239; RECEPTOR\_TYR\_KIN\_II; 1.  
CC PFAM: PF00069; pkinase; 1.  
CC PFAM: PF00360; LRR; 1.  
CC HSP: P11362; IFG1.  
CC TRANSFERASE; TYROSINE-PROTEIN KINASE; ATP-BINDING; PHOSPHORYLATION;  
CC RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT; LEUCINE-REPEAT;  
CC IMMUNOGLOBULIN FOLD; NEUROGENESIS; SIGNAL.  
CC  
CC FT NON\_TER 1 1  
CC FT SIGNAL <1 14  
CC FT CHAIN 15 778  
CC  
CC FT DOMAIN 15 400  
CC FT TRANSMEM 401 421  
CC FT DOMAIN 422 778  
CC FT DOMAIN 51 96  
CC FT REPEAT 51 72  
CC FT REPEAT 75 96  
CC FT REPEAT 193 250  
CC FT DOMAIN 281 347  
CC FT DOMAIN 493 763  
CC FT NP\_BIND 499 507  
CC FT BINDING 527 527  
CC FT ACT\_SITE 633 633  
CC FT MOD\_RES 479 479  
CC FT MOD\_RES 659 659  
CC FT MOD\_RES 663 663  
CC FT MOD\_RES 664 664  
CC FT MOD\_RES 773 773  
CC FT SITE 479 479  
CC  
CC FT SITE 773 773  
CC  
CC FT CARBOHYD 100 100  
CC FT CARBOHYD 130 130  
CC FT CARBOHYD 143 143  
CC FT CARBOHYD 151 151  
CC FT CARBOHYD 194 194  
CC FT CARBOHYD 234 234  
CC FT CARBOHYD 262 262  
CC FT CARBOHYD 300 300  
CC FT CARBOHYD 320 320  
CC FT CARBOHYD 340 340  
CC FT CARBOHYD 384 384  
CC SEQUENCE 778 AA; 87338 MW; 9D26DBD0 CRC32;

Query Match 13.1%; Score 644.5; DB.1; Length 778;  
Best Local Similarity 31.4%; Pred. No. 3e-26;  
Matches 195; Conservative 75; Mismatches 177; Indels 175; Gaps 23;  
  
QY 356 WLLFSEISFSDVNNSSPALGCTFPAPWPGPPPTNFSSLELEPRGQOPVAKPEGSP 415  
DB 226 WELVLEINNISSLNHK-----DLTCRAENSVGLAEDS- 259  
QY 416 TAILIGCLVAIIILLIILALMLRLHWR-----RLLSKARRVLEELTVHLSV- 466  
DB 259 -VMLNVTFFPVILLLSAIPQHEWCFPSVDNPTFRILWLFNGSMLPEGPIHTRIVEY 317  
QY 466 -PGDTILIN-----NRGPREPPPYQEPGRGNPHGAPCPNGS-ALLSNPAYRLLLAT 519  
DB 318 EPNSTVLHGCLQNR-----PTH-----VNNGNYTLVQNP-----LGR 351  
QY 520 YAREPRG-----PGPPTPAWAKPTQAYSGDYMEPEKPGAPLPPPPQNSVPHY 569  
DB 352 AARSIOGRFMDNPFSPPEPIPVISPLGTRNSS-----LEGPVETADEHT 398  
QY 570 AEADIVTLQGV-----TGGNTYAVPALPPG-----AYGDGP 600  
DB 399 FGVSVAVALAFASFLSVMLIALNKGHSKFGINSRAVLAPEDGLAMSLHFWTLGSSP 458  
QY 601 -----PRVD-----PPRSRLRFEKLGEGQGEVHLCVDS--PQ 633  
DB 459 VSSPESKLDGLKSNFIENPOYFCNACVHHVQRDRDVLKWELEGAGFKVFLAECSHLLPE 518  
QY 634 DLVSLDFPLNVRGHPLLVAVKILRPDATKARNDFLKEVKIMSRKLDKPNILLLGVCVQ 693  
DB 519 Q-----EKTLLVAVKALK-EVTENARLDFQREAEELTVLQHEHIVAFYGVCTE 564  
QY 694 DDPLCMITDYMENGLNQFLSAHOLEDKAAEGAPGQQAAGPTISYPMLLHVAQAQIAG 753  
DB 565 GDPLIMVFEYMKHGDNLNRLSHGPDAKILD--QGQPCGQGLTSLH--MLQATQIASG 620  
QY 754 MRYLATLNFVRHDLATRNCLVNGENFTIKIADFGMSRNLVAGDYRYVQGRAVLPIRWMAWE 813  
DB 621 MVLASLHFHVRDLATRNCLVNGHDLVVKIGDFGMSRDIYSTDIYRVGGRMTLPIRWMPPE 680  
QY 814 CILMGKFTTASDWAFVGTLLWEVLMCLCRAOPFGOLTDEQVLENAGFEFFRQGHQVYLSRP 873  
DB 681 SILYRKFTESDIWSFGVWLWEIFYTK-OPWYQLSNTAEIECI-----TQGRE--LRR 732  
QY 874 PACPGQLYELMLRCWSRESEOR 895  
DB 733 RTCESEYDINQSCWQREPOOR 754  
  
RESULT 13  
TRKC\_PIG STANDARD; PRT; 825 AA.  
ID TRKC\_PIG  
AC P24786;  
DT 01-MAR-1992 (REL. 21, CREATED)  
DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)  
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
DE NT-3 GROWTH FACTOR RECEPTOR PRECURSOR (EC 2.7.1.112) (TRKC TYROSINE  
DE KINASE) (GPI45-TRKC) (TRK-C).  
GN NTRK3 OR TRKC.  
OS SUS SCROFA (FIG).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC ARTIODACTYLA; SUIFORMES; SUINA; SUIDAE; SUS.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-BRAIN;  
RX MEDLINE; 91364178.  
RA LAMBALLE F., KLEIN R., BARBACID M.;  
RT "trkc, a new member of the trk family of tyrosine protein kinases, is  
RL a receptor for neurotrophin-3.";  
CC CELL 66:967-979(1991).  
CC -!- FUNCTION: RECEPTOR FOR NEUROTROPHIN-3 (NT-3). THIS IS A TYROSINE-  
CC PROTEIN KINASE RECEPTOR. KNOWN SUBSTRATES FOR THE TRK RECEPTORS  
CC ARE SHC, PI-3 KINASE, AND PLC-GAMMA-1.





15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 NT-3 GROWTH FACTOR RECEPTOR PRECURSOR (EC 2.7.1.112) (TRKC TYROSINE KINASE) (TRK-C).  
 GN. TRK.  
 OS. GALLUS GALLUS (CHICKEN).  
 OC. EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;  
 OC. NEOGNATHAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.  
 RN. [1]  
 RP. SEQUENCE FROM N.A.  
 RX. MEDLINE: 94338700.  
 RA. GARNER A.S., LARGE T.H.;  
 RT. "Isoforms of the avian trkC receptor: a novel kinase insertion  
 RT. dissociates transformation and process outgrowth from survival."  
 RL. NEURON 13:457-472(1994).  
 RN. [2]  
 RP. SEQUENCE OF 378-513 FROM N.A.  
 RX. TISSUE-EMBRYO;  
 RA. MEDLINE: 94084905.  
 RT. WILLIAMS R., BACKSTROM A., EBENDAL T., HALLBOOK F.;  
 RT. "Molecular cloning and cellular localization of trkC in the chicken  
 RT. embryo."  
 RL. BRAIN RES. 75:235-252(1993).  
 CC. -1- FUNCTION: RECEPTOR FOR NEUROTROPHIN-3 (NT-3). THIS IS A TYROSINE-  
 CC. PROTEIN KINASE RECEPTOR. KNOWN SUBSTRATES FOR THE TRK RECEPTORS  
 CC. ARE SHC, PI-3 KINASE, AND PLC-GAMMA-1.  
 CC. -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +  
 CC. PROTEIN TYROSINE PHOSPHATE.  
 CC. -1- PTM: LIGAND-MEDIATED AUTO-PHOSPHORYLATION.  
 CC. -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC. -1- SUBUNIT: EXISTS IN A DYNAMIC EQUILIBRIUM BETWEEN MONOMERIC (LOW  
 CC. AFFINITY) AND DIMERIC (HIGH AFFINITY) STRUCTURES.  
 CC. -1- DEVELOPMENTAL STAGE: EXPRESSION OCCURS IN THE EMBRYONAL DAY 2 (E2)  
 CC. EMBRYO WITH INCREASING LEVELS LATER IN DEVELOPMENT. IN THE E9  
 CC. EMBRYO HIGHEST LEVELS ARE FOUND IN BRAIN AND SPINAL CORD WITH  
 CC. INTERMEDIATE LEVELS IN EYE, HEART, GUT AND MUSCLE. LOW LEVELS ARE  
 CC. FOUND IN KIDNEY, LIVER, SKIN AND YOLK SAC.  
 CC. -1- ALTERNATIVE PRODUCTS: A TRUNCATED FORM IS PRODUCED BY ALTERNATIVE  
 CC. SPLICING. THIS FORM IS PREFERENTIALLY EXPRESSED IN THE ADULT  
 CC. CENTRAL NERVOUS SYSTEM, BUT SCARCELY DURING THE EMBRYONIC STAGES.  
 CC. -1- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-  
 CC. PROTEIN KINASES. HIGHEST SIMILARITY WITH OTHER TRK-TYPE RECEPTORS.  
 CC. -1- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CONTAINS  
 CC. TWO C2-LIKE DOMAIN.  
 CC. -----  
 CC. This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC. between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC. the European Bioinformatics Institute. There are no restrictions on its  
 CC. use by non-profit institutions as long as its content is in no way  
 CC. modified and this statement is not removed. Usage by and for commercial  
 CC. entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC. or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC. -----  
 DR. EMBL: S74248; G712821;  
 DR. EMBL: Z30091; E53026;  
 DR. PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR. PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR. PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR. PROSITE: PS00239; RECEPTOR\_TYR\_KIN\_II; 1.  
 DR. PFAM: PF00047; Ig; 1.  
 DR. PFAM: PF00069; Pkinase; 1.  
 DR. PFAM: PF00560; LRR; 1.  
 DR. HSP: P11362; IFGI.  
 KW. PHOSPHORYLATION: RECEPTOR; GLYCOPROTEIN; NEUROGENESIS; SIGNAL;  
 KW. ALTERNATIVE SPLICING; IMMUNOGLOBULIN FOLD.  
 FT. SIGNAL  
 FT. CHAIN 1 \* 31 BY SIMILARITY.  
 FT. DOMAIN 32 852 NT-3 GROWTH FACTOR RECEPTOR.  
 FT. DOMAIN 32 430 EXTRACELLULAR (POTENTIAL).  
 FT. TRANSMEM 431 455 POTENTIAL.  
 FT. DOMAIN 456 852 CYTOPLASMIC (POTENTIAL).  
 FT. DOMAIN 227 288 IG-LIKE C2-TYPE DOMAIN.  
 FT. DOMAIN 219 382 IG-LIKE C2-TYPE DOMAIN.  
 FT. DOMAIN 340 841 PROTEIN KINASE.

FT NP_BIND	546	554	ATP (BY SIMILARITY).
FT BINDING	574	574	ATP (BY SIMILARITY).
FT ACT_SITE	681	681	BY SIMILARITY.
FT MOD_RES	518	518	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES	707	707	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES	711	711	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES	712	712	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT MOD_RES	847	847	PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT SITE	518	518	INTERACTION WITH SHC PROTEIN (BY SIMILARITY).
FT SITE	847	847	INTERACTION WITH PLC-GAMMA1 (BY SIMILARITY).
FT CARBOHYD	68	68	POTENTIAL.
FT CARBOHYD	72	72	POTENTIAL.
FT CARBOHYD	79	79	POTENTIAL.
FT CARBOHYD	163	163	POTENTIAL.
FT CARBOHYD	203	203	POTENTIAL.
FT CARBOHYD	218	218	POTENTIAL.
FT CARBOHYD	232	232	POTENTIAL.
FT CARBOHYD	259	259	POTENTIAL.
FT CARBOHYD	267	267	POTENTIAL.
FT CARBOHYD	272	272	POTENTIAL.
FT CARBOHYD	294	294	POTENTIAL.
FT CARBOHYD	375	375	POTENTIAL.
FT CARBOHYD	388	388	POTENTIAL.
FT CONFLICT	378	378	I -> F (IN REF. 2).
SQ SEQUENCE	852 AA;	95859 MW;	3B69FB1B CRC32;

Query Match 12.5%; Score 614; DB 1; Length 852;

Best Local Similarity 42.1%; Pred. No. 1.2e-24;

Matches 139; Conservative 49; Mismatches 82; Indels 60; Gaps 10;

QY 607 RSLRKFKEKLGSGGFGVHLCEVDSPODLVSLDRPLNVRKGHPLLVAVKILRPDANKAR 666

Db 537 RRDIVLKLRELGEAGKGVFLAECYNLS-----PTNDK---MLVAVKALK-DPTLAAR 584

QY 667 NDFLKEVKIMSLKDPNIRILGVCVQDDPLCMITDYMGNDLQNFSAHQLEDKAAEGA 726

Db 585 KDFQREAEALLTLQHEHIVKFKVCGCGDPLIMVFYMKHGDNLKFLRAH-----G 635

QY 727 PG-----DQAAAGP-TISYPMLLHVAQAISGRMYLATLNFVHRDLATRNCLYGENFTI 780

Db 636 PDAMILVDGPPQAKGELGSLQMLHIAISQASGMVYLASQHFVHRDLATRNCLVGANLLV 695

QY 781 KIADFGMSRLYAGDYR-----VOGRAVLPIRWMAWECI 815

Db 696 KIGDFGMSRDYVSTDIYREGPRPKQLSTAWQRHLAPPAATVGGHTMLPIRWMPESI 755

QY 816 LMKGFTTASDVWAFVGLTWELVLMCLRAQPFQGLTDEQVFNAGFEFFRDQGRQVYLSRPPA 875

Db 756 MYRKFTTESDVWNSFGVILWEIFTYCK-QPWFLSNTVEICI-----TQGR--VLERPRV 807

QY 876 CPQGLYELMLRCWSRESEORPPFSLHREL 905

Db 808 CPKEYDMLGCMOREPOORLNKEIKYL 837

Search completed: November 4, 1999, 01:54:22

Job time: 7597 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 4, 1999, 01:53:54 ; Search time 25.72 Seconds  
(without alignments)  
954.740 Million cell updates/sec

Title: US-08-170-558-8  
Perfect score: 2175  
Sequence: 1 DADMKGHDPKCRYALGMQ.....LELEPRGQPVAKPEGSPTA 399

Scoring table: BLOSUM62

Searched: 201082 seqs, 61543640 residues

Database : SPTREMBL\_10.\*  
1: sp.archaea:\*  
2: sp.bacteria:\*  
3: sp.fungi:\*  
4: sp.human:\*  
5: sp.invertebrate:\*  
6: sp.mammal:\*  
7: sp.mhc:\*  
8: sp.organelle:\*  
9: sp.phage:\*  
10: sp.plant:\*  
11: sp.rodent:\*  
12: sp.virus:\*  
13: sp.vertebrate:\*  
14: sp.unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1988.5	91.4	911	11	O35407
2	391	18.0	737	5	Q17305
3	381.5	17.5	737	5	Q18163
4	349	16.0	791	5	Q19354
5	222	10.2	931	4	O60462
6	222	10.2	909	4	O14820
7	222	10.2	926	4	O14821
8	220	10.1	909	11	O35373
9	220	10.1	926	11	O35374
10	220	10.1	931	11	O35375
11	220	10.1	901	11	O35376
12	220	10.1	906	11	O35377
13	220	10.1	914	11	O35378
14	217	10.0	925	11	O35276
15	211.5	9.7	845	4	O14113
16	204	9.4	923	4	O60461
17	204	9.4	923	4	O14786
18	201.5	9.3	1128	11	O88442
19	198.5	9.1	784	11	O54860
20	198.5	9.1	722	11	O92100
21	192	8.8	216	4	O14286
22	189.5	8.7	2183	11	O88783
23	184.5	8.5	558	5	O76470
24	184	8.5	2343	6	O18806
25	182	8.4	2343	6	O62730
26	179.5	8.3	2224	4	O43737
27	174	8.0	480	11	O35474
28	172.5	7.9	728	6	O97567
29	167	7.7	480	4	O43854

30	167	7.7	224	4	O15537	O15537 homo sapien
31	164.5	7.6	719	11	Q61281	Q61281 mus musculus
32	163	7.5	224	11	Q921L4	Q921L4 mus musculus
33	150	6.9	1384	4	P78357	P78357 homo sapien
34	148.5	6.8	1383	5	Q94887	Q94887 drosophila
35	135	6.2	1385	11	O54991	O54991 mus musculus
36	135	6.2	1381	11	P97846	P97846 rattus norv
37	134.5	6.2	363	6	O77718	O77718 equus caball
38	106.5	4.9	1008	13	Q9YHV5	Q9YHV5 fugu rubrip
39	105.5	4.9	4123	4	O75851	O75851 homo sapien
40	97	4.5	1101	5	O27235	O27235 drosophila
41	97	4.5	892	5	O27436	O27436 drosophila
42	97	4.5	1269	5	Q05817	Q05817 drosophila
43	94.5	4.3	395	2	P74597	P74597 synechocyst
44	93	4.3	365	2	O927A9	O927A9 chlamydia p
45	92.5	4.3	533	11	Q9Z222	Q9Z222 mus musculus

## ALIGNMENTS

RESULT 1  
O35407  
ID O35407 PRELIMINARY; PRT; 911 AA.  
AC O35407  
DT 01-JAN-1998 (TRENBLrel. 05, Created)  
DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)  
DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
DE RECEPTOR-LIKE TYROSINE KINASE.  
GN NEP.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129;  
RX MEDLINE; 93390947.  
RA ZERLIN M., JULIUS M.A., GOLDFARB M.;  
RT "NEP: a novel receptor-like tyrosine kinase expressed in  
proliferating neuroepithelia";  
RL Oncogene 8:2731-2739(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129;  
RA ZERLIN M., JULIUS M.A., GOLDFARB M.;  
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF026259; AAB81866.1;  
DR PFAM; PF00754; F5\_F8\_type\_C; 1.  
DR PFAM; PF00069; pkinase; 1.  
DR PROSITE; PS00239; RECEPTOR\_TYR\_KIN\_II; 1.  
DR PROSITE; PS01285; FA58C\_1; 1.  
DR PROSITE; PS01286; FA58C\_2; 1.  
SQ SEQUENCE 911 AA; 101091 MW; 6B91AABB CRC32;

Query Match		91.4%	Score 1988.5;	DB 11;	Length 911;
Best Local Similarity		91.8%;	Pred. No. 5.8e-169;		
Matches 367;		Conservative 10;	Mismatches 18;	Indels 5;	Gaps 2;
QY	1	DADMKGHDPKCRYALGMQDRTIPDSISASSWSDDTAARHSRLSSDGDGAWCPAGS	60		
Db	20	DADMKGHDPKCRYALGMQDRTIPDSISASSWSDDTAARHSRLSSDGDGAWCPAGP	79		
QY	61	VFPKKEEYLVQDLQRLHLVALVGTQGRHAGGLGKEFSYRLYSRDGRRWMDKRWGQ	120		
Db	80	VFPKKEEYLVQDLQRLHLVALVGTQGRHAGGLGKEFSYRLYSRDGRRWMDKRWGQ	139		
QY	121	EVISGNEPDEGVVWKDLGPPVRLVREYPRADRVMSVCLRVLYGCLWRDGLLSYAPV	180		
Db	140	EVISGNEPDEGVVWKDLGPPVRLVREYPRADRVMSVCLRVLYGCLWRDGLLSYAPV	199		
QY	181	GOTMYLSEA-VYLDNSTYDGHVTGGLGQLADGVVGLDDFKKSOELRWPGDYVG	239		

Db 200 GQTMLSEVMVHLNDSTYDGTAGGLQYGLGLQADGWGLDDFRSQSELRVWPGYDVG 259  
 Qy 240 WSNHSFSGYVEMFEEDRLRAFQAMQVHCNNHTLGLARLPGGVECFRGRGPAWEGEP 299  
 Db 260 WSNQSFPTGYVEMFEEDRLTFQTMQVHCNNHTLGLARLPGGVECFRGRGPAWEGEP 319  
 Qy 300 MRHLNGLGDPARAVSVPLGGVAVRFLOCRFLFAGPWLFLFSISFISDVWNSSPALG 359  
 Db 320 VRHALLGSLGDPARAVSVPLGGVAVRFLOCRFLFAGPWLFLFSISFISDVWNSSPALG 376  
 Qy 360 GTFFPAPWMPGPPPTNFSSLELEPRGQOPVAKPEGSPTA 399  
 Db 376 DTFPPAPWMPGPPPTNFSSLELEPRGQOPVAKPEGSPTA 415

RESULT 2  
 Q17305 PRELIMINARY; PRT; 737 AA.  
 AC Q17305;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)  
 DE GO1D9.2 PROTEIN.  
 GN GO1D9.2  
 OS Caenorhabditis briggsae.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-GUJARAT G16;  
 RA WU X., LE T.T.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-GUJARAT G16;  
 RA WATERSTON R.;  
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U56248; AAA98703.1;  
 DR PFAM; PF00754; F5\_F8\_type\_C; 1.  
 DR PFAM; PF00069; pkinase.2;  
 SQ SEQUENCE 737 AA; 82950 MW; 36328080 CRC32;

Query Match 18.0%; Score 391; DB 5; Length 737;  
 Best Local Similarity 30.5%; Pred. No. 1e-26;  
 Matches 106; Conservative 60; Mismatches 149; Indels 32; Gaps 14;  
 QY 19 MODRTIPDSDISASSW-SDSTAARHSRLSSDGDGAWCPAGSVFPKKEEYLOVDLQRLH 77  
 Db 1 MONGDVLDAQITASSSFDKQSGVGPONARLHSELASGAWCPKQINSKSYEFLOVTLNDF 60  
 QY 78 LVALVGTQGRHAGGLGKFSRSYRLRYSDGRRWGKWRGQVEISGNEDEPGVVLKDL 137  
 Db 61 LITSVETQGRYNGTGRFASHYMDYMRPGSQWIRYLNRSRSHVYMDGNFTTTPVIRAL 120  
 QY 138 GPMWARLVRFYPRADRVMSVCLRVLYGCLWRDGLSY-TAPVGTMYLSEAVYLNDS 196  
 Db 121 DPPIVASRIKRVFSSKNTRTVCNRAIEHGCK-HEGVYISTVPDGSRL---DTLDFKDSM 176  
 QY 197 YDG---HTVGGGL-QYGGGLQADGVVGLDDFRKQSELRVWPGYDYVWNSHFSFGYVE 251  
 Db 179 NSQMY---TESGIT-RGLGLLTDGYVAQTSPEKQNNNSW-----IGW-NRDTTDRGT 228  
 QY 252 MEFEEDRLRAFQAMQVHCNNHTLGLARLPGGVECFRGRG---PAMAVEGEPMRNLGN 307  
 Db 229 ILFEFEIHNFTDVV-----LATFGNRI-DGIDVIFSDQKTFPLFSQISSSERGLTANT 282  
 QY 308 LGDPRARAVSVPLGGVAVRFLOCRFLFAGPWLFLFSISFISDVWNS 354  
 Db 283 ---SRRYDFRVPLNRAGRKVRISIKFSFSDWMFLTEVHTSGIFKHT 326

RESULT 3  
 Q17305 PRELIMINARY; PRT; 737 AA.

Q18163 PRELIMINARY; PRT; 737 AA.  
 AC Q18163;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)  
 DE SIMILAR TO PROTEIN TYROSINE KINASE.  
 GN C25F6.4  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;  
 OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 94150718.  
 RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
 RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
 RA SMAILDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans.";  
 RL Nature 368:32-38(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA BENILEY D.;  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA WATERSTON R.;  
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U39742; AAA80432.1;  
 DR PFAM; PF00754; F5\_F8\_type\_C; 1.  
 DR PFAM; PF00069; pkinase; 2.  
 SQ SEQUENCE 737 AA; 83396 MW; B5F5CD5A CRC32;

Query Match 17.5%; Score 381.5; DB 5; Length 737;  
 Best Local Similarity 30.3%; Pred. No. 7.3e-26;  
 Matches 103; Conservative 58; Mismatches 150; Indels 29; Gaps 12;  
 QY 19 MODRTIPDSDISASSW-SDSTAARHSRLSSDGDGAWCPAGSVFPKKEEYLOVDLQRLH 77  
 Db 1 MONGDIADSOITASSSFDKQSGVGPONARLHSELASGAWCPKQINSKSYEFLOVTLNDF 60  
 QY 78 LVALVGTQGRHAGGLGKFSRSYRLRYSDGRRWGKWRGQVEISGNEDEPGVVLKDL 137  
 Db 61 LITSVETQGRYNGTGRFASHYMDYLRPGSQWIRYKNRTGHVYMDGNFTTTPVIRVL 120  
 QY 138 GPMWARLVRFYPRADRVMSVCLRVLYGCLWRDGLSY-TAPVGTMYLSEAVYLNDS 196  
 Db 121 DPPIVASRIKRVFSSKNTRTVCNRAIEHGCK-HEGVYISTVPDGSRL---DTLDFKDSM 176  
 QY 197 YDG---HTVGGGL-QYGGGLQADGVVGLDDFRKQSELRVWPGYDYVWNSHFSFGYVE 252  
 Db 177 FEDSQIYTESGIRKRYGSPANRRFRSAASPFEKQNNNSW-----IGW-NRDTTDRITI 230  
 QY 253 EFEDRLRAFQAMQVHCNNHTLGLARLPGGVECFRGRG---PAMAVEGEPMRNLGN 308  
 Db 231 LFEFEVHNFTDVV-----LATFGNRI-DGIDVIFSDQKTFPLFSQISSSERQSLNNT- 284  
 QY 309 GDPRARAVSVPLGGVAVRFLOCRFLFAGPWLFLFSISFISDVWNS 354  
 Db 284 ---SRRYDFRVPLNRAGRKVRISIKFSFSDWMFLTEVHTS 321

RESULT 4  
 Q19354 PRELIMINARY; PRT; 791 AA.  
 ID Q19354



[illegible][illegible]

RESULT	6
014820	PRELIMINARY; PRT; 909 AA.
ID	014820
IAC	014820;
DT	01-JAN-1998 (T=EMBLrel. 05, Created)
DDT	01-JAN-1998 (T=EMBLrel. 05, Last sequence update)
DDT	01-MAY-1999 (T=EMBLrel. 10, Last annotation update)
DE	NEUROFILIN-2.
SOC	Homo sapiens (Human).
OOS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OCN	Eutheria; Primates; Catarrhini; Homnidae; Homo.
CC	[1]
RPR	SEQUENCE FROM N.A.
RRP	MEDLINE; 97470888.
EX	CHEN H.; CHEDOTAL A., HE Z.-G., GOODMAN C.S., TESSIER-LAVIGNE M.;
RA	"Neophilin-2, a novel member of the neuropilin family, is a high
RT	affinity receptor for the semaphorins Sema E and Sema IV but not Sema
RT	III.";
RL	Neuron 19:547-559(1997).
REL	EMBL; AF022859; AAC51788.1; -
DRR	PFAM; PF00431; CUB; 2.
DRD	PFAM; PF00754; F5_F8_type_C; 2.
DRD	PFAM; PF00629; MAW; 1.
DSQ	SEQUENCE 909 AA; 102213 MW; ASF7D0CD CRC32;

Query Match 10.2%; Score 222; DB 4; Length 909;  
Best Local Similarity 32.2%; Pred. No. 1.6e-11;  
Matches 65; Conservative 28; Mismatches 83; Indels 26; Gaps 8;





```

Query Match      10.1%; Score 220; DB 11; Length 906;
Best Local Similarity 31.9%; Pred. No. 2.4e-11;
Matches 59; Conservative 31; Mismatches 69; Indels 26; Gaps 8;

QY 7 HFDPA---KCRYALGMDRTIPDSISASSWSGST-AARHSRLSSDGDGAWCPAGSVF 62
Db HOEPPENFQCNVPLGMSGRIANEIOISASTFSDGRWTPQOSRLHGD--DNGWTPN---L 322
QY 63 PKEEYLQVDLQRLHLVALVGTQGRHAGLGKE-----FSRSYRLRYSRDGRRWGWKDR 117
Db 323 DSNKEYLQVDLRLTMTAIAATQ---GAISRETQKGYVVKSYKLEVTNGEDMWVYRHG 378
QY 118 WQGEVISGNEDEGVVLKDLGPPVAVLRVFPYPRADRVMSVCLRVLYGCLWRD----- 172
Db 379 KNHKIFQANNDATVVLNKLHMLPLTRFIRIPQTHW-LGIALRLLEFGCRVTDAPCSNM 437

QY 172 -GLLS 175
Db 438 LGMLS 442

RESULT 13
O35378
ID O35378 PRELIMINARY; PRT; 914 AA.
AC O35378;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE NEUROPILIN 2.
GN NR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RC STRAIN-BALB/C;
RX CHEN H., CHEDOTAL A., HE Z.-G., GOODMAN C.S., TESSIER-LAVIGNE M.;
RA "Neuropilin-2, a novel member of the neuropilin family, is a high
RT affinity receptor for the semaphorins Sema E and Sema IV but not Sema
RT III."
RL Neuron 19:547-559(1997).
DR EMBL; AF022861; AAC53382.1; -.
DR MGD; MGI:1100492; NRP2.
DR PFAM; PF00431; CUB; 2.
DR PFAM; PF00754; F5_F8_type_C; 2.
DR PFAM; PF00629; NAM; 1.
SQ SEQUENCE 914 AA; 102701 MW; 5BE1E858 CRC32;

Query Match      10.1%; Score 220; DB 11; Length 914;
Best Local Similarity 31.9%; Pred. No. 2.4e-11;
Matches 59; Conservative 31; Mismatches 69; Indels 26; Gaps 8;

QY 7 HFDPA---KCRYALGMDRTIPDSISASSWSGST-AARHSRLSSDGDGAWCPAGSVF 62
Db HOEPPENFQCNVPLGMSGRIANEIOISASTFSDGRWTPQOSRLHGD--DNGWTPN---L 322
QY 63 PKEEYLQVDLQRLHLVALVGTQGRHAGLGKE-----FSRSYRLRYSRDGRRWGWKDR 117
Db 323 DSNKEYLQVDLRLTMTAIAATQ---GAISRETQKGYVVKSYKLEVTNGEDMWVYRHG 378
QY 118 WQGEVISGNEDEGVVLKDLGPPVAVLRVFPYPRADRVMSVCLRVLYGCLWRD----- 172
Db 379 KNHKIFQANNDATVVLNKLHMLPLTRFIRIPQTHW-LGIALRLLEFGCRVTDAPCSNM 437

QY 172 -GLLS 175
Db 438 LGMLS 442

RESULT 14
O35378
ID O35378 PRELIMINARY; PRT; 914 AA.
AC O35378;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE NEUROPILIN 2.
GN NR2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN [1]
RC STRAIN-BALB/C;
RX CHEN H., CHEDOTAL A., HE Z.-G., GOODMAN C.S., TESSIER-LAVIGNE M.;
RA "Neuropilin-2, a novel member of the neuropilin family, is a high
RT affinity receptor for the semaphorins Sema E and Sema IV but not Sema
RT III."
RL Neuron 19:547-559(1997).
DR EMBL; AF022861; AAC53382.1; -.
DR MGD; MGI:1100492; NRP2.
DR PFAM; PF00431; CUB; 2.
DR PFAM; PF00754; F5_F8_type_C; 2.
DR PFAM; PF00629; NAM; 1.
SQ SEQUENCE 914 AA; 102701 MW; 5BE1E858 CRC32;

```

```

O35276
ID O35276 PRELIMINARY; PRT; 925 AA.
AC O35276;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE NEUROPILIN-2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RC STRAIN-SPRAGUE-DAWLEY;
RX MEDLINE; 97433085.
RA KOLODKIN A.L., LEVENGOOD D.V., ROWE E.G., TAI Y.-T., GIGER R.J.;
RA GINTY D.D.;
RT "Neuropilin is a semaphorin III receptor."
RL Cell 90:753-762(1997).
DR EMBL; AF016297; AAC53338.1; -.
DR PFAM; PF00431; CUB; 2.
DR PFAM; PF00754; F5_F8_type_C; 2.
DR PFAM; PF00629; NAM; 1.
SQ SEQUENCE 925 AA; 103896 MW; B097CCCCA CRC32;

Query Match      10.0%; Score 217; DB 11; Length 925;
Best Local Similarity 31.9%; Pred. No. 4.5e-11;
Matches 59; Conservative 31; Mismatches 69; Indels 26; Gaps 8;

QY 7 HFDPA---KCRYALGMDRTIPDSISASSWSGST-AARHSRLSSDGDGAWCPAGSVF 62
Db HOEPPENFQCNVPLGMSGRIANEIOISASTFSDGRWTPQOSRLHGD--DNGWTPN---V 322
QY 63 PKEEYLQVDLQRLHLVALVGTQGRHAGLGKE-----FSRSYRLRYSRDGRRWGWKDR 117
Db 323 DSNKEYLQVDLRLTMTAIAATQ---GAISRETQKGYVVKSYKLEVTNGEDMWVYRHG 378
QY 118 WQGEVISGNEDEGVVLKDLGPPVAVLRVFPYPRADRVMSVCLRVLYGCLWRD----- 172
Db 379 KNHKIFQANNDATVVLNKLHMLPLTRFIRIPQTHW-LGIALRLLEFGCRVTDAPCSNM 437

QY 172 -GLLS 175
Db 438 LGMLS 442

RESULT 15
O35276
ID O35276 PRELIMINARY; PRT; 845 AA.
AC O35276;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE AEBPL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominiidae; Homo.
RN [1]
RC TISSUE=CANCELLOUS BONE;
RX MEDLINE; 97079196.
RA OHNO I., HASHIMOTO J., SHIMIZU K., TAKAOKA K., OCHI T., MATSUBARA K.;
RA OKUBO K.;
RT "A cDNA cloning of human AEBPL from primary cultured osteoblasts and
RT its expression in a differentiating osteoblastic cell line."
RL Biochem. Biophys. Res. Commun. 228:411-414(1996).
DR EMBL; D86479; BAAL3094.1; -.
DR PFAM; PF00754; F5_F8_type_C; 1.
DR PFAM; PF00246; Zn_carboxypept; 2.
SQ SEQUENCE 845 AA; 96173 MW; 81C8A087 CRC32;

Query Match      9.7%; Score 211.5; DB 4; Length 845;

```

Best Local Similarity 24.38; Pred. No. 1.2e-10;  
Matches 80; Conservative 54; Mismatches 144; Indels 51; Gaps 12;

Qy	9	DPACRYALGMDRTIPDSISASSSSWSDSTAARHSRLSSDG-----DGAWCPAGSV	61
Db	68	EKVKCP-PIGNESHRIEDNOIRASSMLRHGLGARGELNMQTGATEDDDYDGAACAEEDA	126
Qy	62	FPKEEYLQVDLQRLHLVALVGTQGRHAGGLGKEFSYRLRYSRDGRRMGMKDRWGOE	121
Db	127	---RTQWIEVDTRRTFTGVTITQGRDS-SIHDDFVTFVFGFSNDSQTWVMYTNGEEM	182
Qy	122	VISGNEDEPGVWLKLGPPVVARLVREYPRADRVMSVCLVELYGCILWRDGLLSYTPVG	181
Db	183	TFHGNVDKDTPVUSELPEPVVARFIRIYPLTN-GLCMRLEVLC-----SVAPY	233
Qy	182	QTYLSEAVYLNDSTYDGTGVGLQYGGGLQGLADGV-----VGLDDFRKSQEL	229
Db	234	SYAAQNEVATDDLDFRHS-----YKDMRQLMKVYNEECPTITRTYSLG----KSRGL	284
Qy	230	RVW-----PGDYVGNWSNHSFSSGYVMEFEFDRLRAFQAMQVHCNNMHTLGARLPG	282
Db	285	KIYAMEISDNPGHEHGEPEFRYTAGIHGNEV-LGRELLLLLMQYLCREYRDGNPRVRS	343
Qy	283	V-ECFRFRGPAMAWGEPMRHNLGGNLGD	310
Db	344	VQDTRIHLVPSLNDPDGYEVAQAQMGSEFGN	372

Search completed: November 4, 1999, 01:53:56  
Job time: 7616 sec



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 4, 1999, 01:54:22 ; Search time 15.86 Seconds  
(without alignments)  
711.163 Million cell updates/sec

Title: US-08-170-558-8  
Perfect score: 2175  
Sequence: 1 DADMKGHDPKCRYALGMQ.....LELEPRGQVPVAKPEGSPTA 399

Scoring table: BLOSUM62

Searched: 77977 seqs, 28268293 residues

Database : SwissProt\_37:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2167	99.6	913	1	EDD1_HUMAN	Q08345 homo sapien
2	2021	92.9	910	1	EDD1_RAT	Q63474 rattus norv
3	1999.5	91.9	911	1	EDD1_MOUSE	Q03146 mus musculu
4	1002.5	46.1	855	1	TRK3_HUMAN	Q16832 homo sapien
5	995.5	45.8	854	1	TRK3_MOUSE	Q62371 mus musculu
6	213	9.8	928	1	NRP_XENLA	P28824 xenopus lae
7	210	9.7	2133	1	FA8_PIG	P12263 sus scrofa
8	204	9.4	923	1	NRP_MOUSE	P97333 mus musculu
9	192	8.8	2351	1	FA8_HUMAN	P00451 homo sapien
10	189	8.7	914	1	NRP_CHICK	P79795 gallus gall
11	178.5	8.2	2224	1	FA5_HUMAN	P12259 homo sapien
12	177.5	8.2	387	1	MFGM_HUMAN	Q08431 homo sapien
13	175	8.0	2319	1	FA8_MOUSE	Q06194 mus musculu
14	174.5	8.0	2211	1	FA5_BOVIN	Q28107 bos taurus
15	174.5	8.0	463	1	MFGM_MOUSE	P21956 mus musculu
16	172.5	7.9	427	1	MFGM_RAT	P70490 rattus norv
17	167.5	7.7	427	1	MFGM_BOVIN	Q95114 bos taurus
18	166.5	7.7	409	1	MFGM_PIG	P79385 sus scrofa
19	145.5	6.7	3133	1	HMCT_BOMMO	P98092 bombyx mori
20	143	6.6	257	1	DIS2_DICDI	P42530 dictyosteli
21	106.5	4.9	448	1	HGD_LEMENI	Q00667 emericella
22	97	4.5	1101	1	FURL_DROME	P30430 drosophila
23	97	4.5	899	1	FURS_DROME	P26016 drosophila
24	97	4.5	789	1	LP1B_DROME	P11996 drosophila
25	96	4.4	253	1	DISC_DICDI	P02887 dictyosteli
26	92.5	4.3	985	1	AGLU_ASPNG	P56526 aspergillus
27	92	4.2	635	1	SIR_SIRNY3	P72854 synchocyst
28	91.5	4.2	411	1	CYB_CHRVI	O31215 chromatiu
29	91	4.2	253	1	DISA_DICDI	P02886 dictyosteli
30	89	4.1	477	1	CSCA_ECOLI	P40714 escherichia
31	89	4.1	867	1	SSPO_BOVIN	P98167 bos taurus
32	89	4.1	3164	1	TEGU_HSV11	P10220 herpes simp
33	87.5	4.0	139	1	LAMP_PETMA	P33575 petromyzon
34	87.5	4.0	1168	1	MYSC_ACACA	P10569 acanthamoeb
35	87	4.0	439	1	HFH1_RAT	Q63244 rattus norv
36	86.5	4.0	149	1	DISD_DICDI	P02888 dictyosteli
37	86.5	4.0	379	1	UL28_HCMVA	P16847 human cytom
38	86	4.0	253	1	RM02_ACACA	P46763 acanthamoeb
39	85.5	3.9	3567	1	ERY2_SACER	Q03132 saccharopol
40	85.5	3.9	1460	1	RPC1_YEAST	P04051 saccharomyc
41	84.5	3.9	393	1	CYB_METSE	O47499 metridium s
42	84	3.9	548	1	FLR1_YEAST	P38124 saccharomyc
43	84	3.9	416	1	HXD3_HUMAN	P31249 homo sapien

#### ALIGNMENTS

```
RESULT 1
ID EDD1_HUMAN STANDARD; PRT: 913 AA.
AC Q08345; Q14196; Q16562;
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE EPIHELIAL DISCOLDIN DOMAIN RECEPTOR 1 PRECURSOR (EC 2.7.1.112)
DE (TYROSINE-PROTEIN KINASE CAK) (CELL ADHESION KINASE) (TYROSINE KINASE
DE (DISCOLDIN RECEPTOR TYROSINE KINASE) (TRK E) (PROTEIN-TYROSINE
DE KINASE RTK 6).
GN EDD1 OR CAK OR DDR1 OR TRKE OR RTK6.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FETAL LIVER;
RX MEDLINE; 94134417.
RA PEREZ J L., SHEN X., FINKERNAGEL S., SCIORRA L., JENKINS N.A.,
RA GILBERT D.J., COELAND N.G., WONG T.W.;
RT "Identification and chromosomal mapping of a receptor tyrosine kinase
RT with a putative phospholipid binding sequence in its ectodomain.";
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE; 93296201.
RA JOHNSON J.D., EDMAN J.C., RUTTER W.J.;
RT "A receptor tyrosine kinase found in breast carcinoma cells has an
RT extracellular discoidin I-like domain.";
RN [3]
RP PROC. NATL. ACAD. SCI. U.S.A. 90:5677-5681(1993).
RX MEDLINE; 97131588.
RA SAKUMA S., TADA M., SAYA H., SAWAMURA Y., SHINOHE Y., ABE H.;
RT "Receptor protein tyrosine kinase DDR is up-regulated by p53
RT protein.";
RN [4]
RP FEBS LETT. 398:165-169(1996).
RX SEQUENCE FROM N.A.
RC TISSUE=FROM N.A. (SHORT FORM).
RX MEDLINE; 96389017.
RA PLAYFORD M.P., BUTLER R.J., WANG X.C., KATSO R.M., COOKE I.E.,
RA GANESAN T.S.;
RT "The genomic structure of discoidin receptor tyrosine kinase.";
RN [5]
RP GENOME RES. 6:620-627(1996).
RX SEQUENCE FROM N.A. (SHORT FORM).
RC TISSUE=LUNG;
RX MEDLINE; 96204002.
RA PEREZ J L., JING S.O., WONG T.W.;
RT "Identification of two isoforms of the Cak receptor kinase that are
RT coexpressed in breast tumor cell lines.";
RN [6]
RP ONCOGENE 12:1469-1477(1996).
RX SEQUENCE FROM N.A. (SHORT FORM).
RC TISSUE=BRAIN, AND KERATINOCYTES;
RX MEDLINE; 94043265.
RA DI MARCO E., CUTULI N., GUERRA L., CANCEDDA R., DE LUCA M.;
RT "Molecular cloning of trke, a novel trk-related putative tyrosine
RT kinase receptor isolated from normal human keratinocytes and widely
RT expressed by normal human tissues.";
RN [7]
RP J. BIOL. CHEM. 268:24290-24295(1993).
RX SEQUENCE FROM N.A. (SHORT FORM).
RC TISSUE-OVARY;
```

44 83.5 3.8 362 1 LOKO GORGO P30388 gorilla gor  
45 --82.5 3.8 331 1 PRP1\_HUMAN P04280 homo sapien

RX MEDLINE: 95151638.  
 RA LAVAL S., BUTLER R., SHELLING A.N., HANBY A.M., POULSOM R.,  
 RT GANESAN T.S.:  
 "Isolation and characterization of an epithelial-specific receptor  
 tyrosine kinase from an ovarian cancer cell line."  
 RL CELL GROWTH DIFFER. 5:1173-1183(1994).  
 CC -!- FUNCTION: MAY BE INVOLVED IN CELL-CELL INTERACTIONS AND  
 RECOGNITION.  
 CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE -> ADP +  
 PROTEIN TYROSINE PHOSPHATE.  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -!- TISSUE SPECIFICITY: EXPRESSED AT LOW LEVELS IN MOST ADULT  
 TISSUES AND IS HIGHEST IN THE BRAIN AND LUNG. ABUNDANT IN  
 BREAST CARCINOMA CELL LINES.  
 CC -!- DOMAIN: THE GLY/PRO-RICH DOMAINS MAY BE REQUIRED FOR AN UNUSUAL  
 GEOMETRY OF INTERACTION WITH LIGAND OR SUBSTRATES.  
 CC -!- ALTERNATIVE PRODUCTS: TWO ISOFORMS ARE PRODUCED BY ALTERNATIVE  
 SPLICING. THE LONG FORM (CALLED CAK I AND SHOWN HERE) AND THE  
 SHORT FORM (CALLED CAK II) ONLY DIFFER BY THE ABSENCE OF A 37  
 RESIDUES SEGMENT.  
 CC -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC  
 DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 1 F5/8 TYPE C DOMAIN.  
 CC -!- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-  
 PROTEIN KINASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L11315; G403387; -  
 DR EMBL: L20817; G306475; -  
 DR EMBL: U48705; G1754536; -  
 DR EMBL: X98208; E251231; -  
 DR EMBL: X99023; E251231; JOINED.  
 DR EMBL: X99024; E251231; JOINED.  
 DR EMBL: X99025; E251231; JOINED.  
 DR EMBL: X99026; E251231; JOINED.  
 DR EMBL: X99027; E251231; JOINED.  
 DR EMBL: X99028; E251231; JOINED.  
 DR EMBL: X99029; E251231; JOINED.  
 DR EMBL: X99030; E251231; JOINED.  
 DR EMBL: X99031; E251231; JOINED.  
 DR EMBL: X99032; E251231; JOINED.  
 DR EMBL: X99033; E251231; JOINED.  
 DR EMBL: X99034; E251231; JOINED.  
 DR EMBL: L57508; G1160925; -  
 DR EMBL: X74979; G400463; -  
 DR EMBL: Z29093; G732800; -  
 DR MIM: 600432; -  
 DR MIM: 600408; -  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; FALSE\_NEG.  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE: PS00111; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00239; RECEPTOR\_TYR\_KIN\_II; 1.  
 DR PROSITE: PS01285; FA58C\_1; 1.  
 DR PROSITE: PS01286; FA58C\_2; 1.  
 DR PFAM: PF00754; F5\_F8\_type\_C; 1.  
 DR HSSP: P11362; 1FGI.  
 DR TRANSFERASE: TYROSINE-PROTEIN KINASE; GLYCOPROTEIN; SIGNAL;  
 KW PHOSPHORYLATION; TRANSMEMBRANE; RECEPTOR; ATP-BINDING;  
 KW ALTERNATIVE SPLICING.  
 FT SIGNAL 1 18 POTENTIAL.  
 FT CHAIN 19 913 EPITHELIAL DISCOIDIN DOMAIN RECEPTOR 1.  
 FT DOMAIN 19 416 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 417 443 POTENTIAL.  
 FT DOMAIN 444 913 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 31 185 F5/8 TYPE C (PHOSPHOLIPID-BINDING,

FT DOMAIN 377 415 POTENTIAL).  
 FT DOMAIN 476 601 GLY/PRO-RICH.  
 FT DOMAIN 610 624 GLY/PRO-RICH.  
 FT NP\_BIND 616 624 PROTEIN KINASE.  
 FT BINDING 655 655 ATP (BY SIMILARITY).  
 FT ACT\_SITE 766 766 ATP (BY SIMILARITY).  
 FT DISULFID 31 185 BY SIMILARITY.  
 FT MOD\_RES 513 513 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD\_RES 792 792 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD\_RES 796 796 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD\_RES 797 797 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CARBOHYD 211 211 POTENTIAL.  
 FT CARBOHYD 260 260 POTENTIAL.  
 FT CARBOHYD 371 371 POTENTIAL.  
 FT CARBOHYD 394 394 POTENTIAL.  
 FT CARSPLIC 506 542 MISSING (IN SHORT FORM).  
 FT CONFLICT 94 94 L -> V (IN REF. 2 AND 3).  
 FT CONFLICT 833 833 L -> V (IN REF. 2 AND 3).  
 FT CONFLICT 847 867 QLTDEQVIEANAGEFFRDGGRQ -> SAHRRAGHRRGGVLP  
 GPQA (IN REF. 4).  
 FT SEQUENCE 913 AA: 101127 MW: 16071364 CRC32:  
 Query Match 99.6%; Score 2167; DB 1; Length 913;  
 Best Local Similarity 99.7%; Pred. No. 1.1e-170;  
 Matches 398; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 DADMGHFDPAKCRVAGLGMODRTIPDSISASSWSNSTAAHRSLESDDGAWCPAGS 60  
 DB 19 DADMGHFDPAKCRVAGLGMODRTIPDSISASSWSNSTAAHRSLESDDGAWCPAGS 78  
 QY 61 VFPKEEYLVQDLQRLHLVALVGTQGRHAGGLGKEFSYRLRYSDGRRMMWMDKRWGQ 120  
 DB 79 VFPKEEYLVQDLQRLHLVALVGTQGRHAGGLGKEFSYRLRYSDGRRMMWMDKRWGQ 138  
 QY 121 EVISGNEDEPVGVLKDLGPPMVARLVRFPYPRADRVMSVCLRVLYGCLWRDGLLSYAPV 180  
 DB 139 EVISGNEDEPVGVLKDLGPPMVARLVRFPYPRADRVMSVCLRVLYGCLWRDGLLSYAPV 198  
 QY 181 GQTMYLEAVYLNDSTYDGHVTGGYGLGQADGVGLDFFKRSQELRVWPGYDYVGW 240  
 DB 199 GQTMYLEAVYLNDSTYDGHVTGGYGLGQADGVGLDFFKRSQELRVWPGYDYVGW 258  
 QY 241 SNHSFSGYVMEFEFDRLEAFQAMQVHCHNNMHTLGARLPGGVGECFRFRGPANAVEGEPM 300  
 DB 259 SNHSFSGYVMEFEFDRLEAFQAMQVHCHNNMHTLGARLPGGVGECFRFRGPANAVEGEPM 318  
 QY 301 RHNLGNLGDPRARAVSVPLGGRVAREFLQCLFAGPWLFFSEISFISDVVNNSSPALGG 360  
 DB 319 RHNLGNLGDPRARAVSVPLGGRVAREFLQCLFAGPWLFFSEISFISDVVNNSSPALGG 378  
 QY 361 TFPAPWPPGPPPTNFSSLELEPRGQQQVAKPEGSPTA 399  
 DB 379 TFPAPWPPGPPPTNFSSLELEPRGQQQVAKPEGSPTA 417  
 RESULT 2  
 EDDL\_RAT  
 ID EDDL\_RAT STANDARD; PRT; 910 AA.  
 AC O63474;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE EPITHELIAL DISCOIDIN DOMAIN RECEPTOR 1 PRECURSOR (EC 2.7.1.112)  
 DE (TYROSINE-PROTEIN KINASE CAK) (CELL ADHESION KINASE) (TYROSINE KINASE  
 DE DDR) (DISCOIDIN RECEPTOR TYROSINE KINASE) (PROTEIN-TYROSINE KINASE  
 DE PK-3).  
 GN EDDL OR PK3.  
 OS RATTUS NORVEGICUS (RAT).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; RATTUS.  
 RN [1]



SEQUENCE FROM N.A.  
 STRAIN-SPRAGUE-DAWLEY; TISSUE=BRAIN;  
 MEDLINE; 94173920.  
 RA SANCHEZ M.P., TAPLEY P., SAINI S.S., HE B., PULIDO D., BARBACID M.;  
 RT "Multiple tyrosine protein kinases in rat hippocampal neurons:  
 RT isolation of PK-3, a receptor expressed in proliferative zones of  
 RT the developing brain";  
 RL PROC. NATL. ACAD. SCI. U.S.A. 91:1819-1823(1994).  
 CC -1- FUNCTION: MAY BE INVOLVED IN CELL-CELL INTERACTIONS AND  
 CC RECOGNITION (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +  
 CC PROTEIN TYROSINE PHOSPHATE.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -1- TISSUE SPECIFICITY: VARIOUS EMBRYONIC AND ADULT TISSUES; ALSO  
 CC PROLIFERATIVE ZONES OF THE DEVELOPING BRAIN; HIPPOCAMPAL NEURONS.  
 CC -1- SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF TYR-PROTEIN  
 CC KINASES.  
 CC -1- SIMILARITY: CONTAINS 1 F5/8 TYPE C DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-  
 CC PROTEIN KINASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L26525; G432481;  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS00239; RECEPTOR\_TYR\_KIN\_II; 1.  
 DR PROSITE; PS01285; FA58C\_1; 1.  
 DR PROSITE; PS01286; FA58C\_2; 1.  
 DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
 DR PFAM; PF00069; pkinase; 1.  
 DR PFAM; PF00754; F5\_F8\_type\_C; 1.  
 DR HSP; P11362; 1FG1.  
 KW TRANSFERASE; TYROSINE-PROTEIN KINASE; GLYCOPROTEIN; SIGNAL;  
 KW PHOSPHORYLATION; TRANSMEMBRANE; RECEPTOR; ATP-BINDING.  
 FT SIGNAL 1 19 POTENTIAL.  
 FT CHAIN 20 910 EPIHELIAL DISCOIDIN DOMAIN RECEPTOR 1.  
 FT DOMAIN 20 413 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 414 440 POTENTIAL.  
 FT DOMAIN 441 910 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 32 186 F5/8 TYPE C (PHOSPHOLIPID-BINDING,  
 FT POTENTIAL).  
 FT DOMAIN 378 412 GLY/PRO-RICH.  
 FT DOMAIN 473 598 GLY/PRO-RICH.  
 FT DOMAIN 607 902 PROTEIN KINASE.  
 FT NP\_BIND 613 621 ATP (BY SIMILARITY).  
 FT BINDING 652 652 ATP (BY SIMILARITY).  
 FT ACT\_SITE 763 763 BY SIMILARITY.  
 FT DISULFID 32 186 BY SIMILARITY.  
 FT MOD\_RES 510 510 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD\_RES 789 789 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD\_RES 793 793 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD\_RES 794 794 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CARBOHYD 212 212 POTENTIAL.  
 FT CARBOHYD 261 261 POTENTIAL.  
 FT CARBOHYD 371 371 POTENTIAL.  
 FT CARBOHYD 391 391 POTENTIAL.  
 SQ SEQUENCE 910 AA; 101164 MW; D69094A9 CRC32;

Query Match 92.9%; Score 2021; DB 1; Length 910;  
 Best Local Similarity 93.0%; Pred. No. 1.2e-158;  
 Matches 371; Conservative 9; Mismatches 15; Indels 4; Gaps 1;  
 QY 1 DADMKGHFDPAKRYALGMQDRTPDSDISASSWSWSDTAARSRLESSDGDGAWCPAGS 60  
 DB 20 DADMKGHFDPAKRYALGMQDRTPDSDISASSWSWSDTAARSRLESSDGDGAWCPAGP 79

QY 61 VPKKEEYLQVDLQRLHLVALVGTQGRHAGGLGKEFSRSYRLRYSDGRRRMGWKDRWGQ 120  
 DB 80 VPKKEEYLQVDLQRLHLVALVGTQGRHAGGLGKEFSRSYRLRYSDGRRRMGWKDRWGQ 139  
 QY 121 EVISGNEDEPGVVLKDLGPPMVARLVREYPRADRVMSVCLRVLYGCLWRDGLLSYTA 180  
 DB 140 EVISGNEDEPGVVLKDLGPPMVARLVREYPRADRVMSVCLRVLYGCLWRDGLLSYTA 199  
 QY 181 GQTMWSEAVYLNDSYDGHVGGVGLGGLGOLADGVVGLDDFRKSOELRWPGYDVGW 240  
 DB 200 GQTMWSEAVYLNDSYDGHVGGVGLGGLGOLADGVVGLDDFRKSOELRWPGYDVGW 259  
 QY 241 SNHSFSSGVEVEFEEDRLRAFOAMOVHCNNMHTLGARLPGVCECRFRGPPAMAWEGEP 300  
 DB 260 SNHSFSSGVEVEFEEDRLRAFOAMOVHCNNMHTLGARLPGVCECRFRGPPAMAWEGEP 319  
 QY 301 RNLGNLGDPRARAVSVPLGGRVAREFLQCRFLFAGPWLFLSEISFISDVVNSSPALGG 360  
 DB 320 RNLGNLGDPRARAVSVPLGGRVAREFLQCRFLFAGPWLFLSEISFISDVVNSSPALGG 375  
 QY 361 TFPAPWPPGPPPTNFSSLELEPRGQOPVAKPEGSP 399  
 DB 376 TFPAPWPPGPPPTNFSSLELEPRGQOPVAKPEGSP 414  
 RESULT 3  
 EDDL\_MOUSE  
 ID EDDL\_MOUSE STANDARD; PRT; 911 AA.  
 AC Q03146;  
 DT 01-OCT-1994 (REL. 30, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE EPIHELIAL DISCOIDIN DOMAIN RECEPTOR 1 PRECURSOR (EC 2.7.1.112)  
 DE (TYROSINE-PROTEIN KINASE CAK) (CELL ADHESION KINASE) (TYROSINE KINASE  
 DE DDR) (DISCOIDIN RECEPTOR TYROSINE KINASE) (PROTEIN-TYROSINE KINASE  
 DE MPK-6).  
 GN EDDR1 OR CAK OR MPK6.  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MORINAE; MUS.  
 [1]  
 SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.  
 RN STRAIN-C57BL;  
 RC MEDLINE; 96204002.  
 RA PEREZ J.L., JING S.Q., WONG T.W.;  
 RT "Identification of two isoforms of the Cak receptor kinase that are  
 RT coexpressed in breast tumor cell lines.";  
 RL ONCOGENE 12:1469-1477(1996).  
 [2]  
 SEQUENCE OF 766-822 FROM N.A.  
 RN STRAIN-C57BL; TISSUE=EMBRYONIC BRAIN;  
 RC MEDLINE; 93096484.  
 RA GILARDI-HEBENSTREIT P., NIETO M.A., FRAIN M., MATTEI M.-G.,  
 RA CHESTER A., WILKINSON D.G., CHARNAY P.;  
 RT "An Eph-related receptor protein tyrosine kinase gene segmentally  
 RT expressed in the developing mouse hindbrain.";  
 RL ONCOGENE 7:2499-2506(1992).  
 CC -1- FUNCTION: MAY BE INVOLVED IN CELL-CELL INTERACTIONS AND  
 CC RECOGNITION (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +  
 CC PROTEIN TYROSINE PHOSPHATE.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -1- TISSUE SPECIFICITY: THE PREDOMINANT ISOFORM CAK I, IS EXPRESSED IN  
 CC DEVELOPING EMBRYO AND ADULT BRAIN; CAK II IS EXPRESSED IN VARIOUS  
 CC EPITHELIAL CELLS.  
 CC -1- ALTERNATIVE PRODUCTS: TWO ISOFORMS ARE PRODUCED BY ALTERNATIVE  
 CC SPLICING. THE LONG FORM (CALLED CAK I AND SHOWN HERE) AND THE  
 CC SHORT FORM (CALLED CAK II) ONLY DIFFER BY THE ABSENCE OF A 37  
 CC RESIDUES SEGMENT.  
 CC -1- SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF TYR-PROTEIN  
 CC KINASES.  
 CC -1- SIMILARITY: CONTAINS 1 F5/8 TYPE C DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-

PROTEIN KINASES.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL: L57509; GI161063; -  
EMBL: X57240; G53198; -  
PIR: S30502; S30502.  
MGD: MGI:99216; CAK.  
DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; FALSE\_NEG.  
DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE: PS00111; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE: PS00239; RECEPTOR\_TYR\_KIN\_II; 1.  
DR PROSITE: PS01285; FA58C\_1; 1.  
DR PROSITE: PS01286; FA58C\_2; 1.  
DR PFAM: PF00069; pkinase; 1.  
DR PFAM: PF00754; F5\_F8\_type\_C; 1.  
DR HSP: P11362; 1FGI.  
KW TRANSFERASE: TYROSINE-PROTEIN KINASE; GLYCOPROTEIN; SIGNAL;  
KW PHOSPHORYLATION: TRANSMEMBRANE; RECEPTOR: ATP-BINDING;  
KW ALTERNATIVE SPLICING.  
FT SIGNAL 1 19 POTENTIAL.  
FT CHAIN 20 911 EPITHELIAL DISCOLDIN DOMAIN RECEPTOR 1.  
FT DOMAIN 20 414 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 415 441 POTENTIAL.  
FT DOMAIN 442 911 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 32 186 F5/8 TYPE C (PHOSPHOLIPID-BINDING, POTENTIAL).  
FT DOMAIN 379 413 GLY/PRO-RICH.  
FT DOMAIN 474 599 GLY/PRO-RICH.  
FT DOMAIN 608 903 PROTEIN KINASE.  
FT NP\_BIND 614 622 ATP (BY SIMILARITY).  
FT BINDING 653 764 ATP (BY SIMILARITY).  
FT ACT\_SITE 764 764 BY SIMILARITY.  
FT DISULFID 32 186 BY SIMILARITY.  
FT MOD\_RES 511 511 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
FT MOD\_RES 790 790 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
FT MOD\_RES 794 794 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
FT MOD\_RES 795 795 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
FT CARBOHYD 213 213 POTENTIAL.  
FT CARBOHYD 262 262 POTENTIAL.  
FT CARBOHYD 372 372 POTENTIAL.  
FT CARBOHYD 392 392 POTENTIAL.  
FT VARSPLIC 503 539 MISSING (IN VARIANT CAK II).  
SQ SEQUENCE 911 AA; 101160 MW; 3DCBB321 CRC32;

Query Match 91.9%; Score 1999.5; DB 1; Length 911;  
Best Local Similarity 92.2%; Pred. No. 6.8e-157; Indels 5; Gaps 2;  
Matches 369; Conservative 10; Mismatches 16.

QY 1 DADMKGHFPAKRYALGMODRIPDSISASSSSWSDSTAARHSRLSSDGDGCWCPAGS 60  
DB 20 DADMKGHFPAKRYALGMODRIPDSISASSSSWSDSTAARHSRLSSDGDGCWCPAGP 79  
QY 61 VFPKEEYLQVDLQRLHLVALVTQGRHAGGLGKEFSRSLRYSDRGRRWGWKDRWQ 120  
DB 80 VFPKEEYLQVDLQRLHLVALVTQGRHAGGLGKEFSRSLRYSDRGRRWGWKDRWQ 139  
QY 121 EVISGNEDPEGVVLKDLGPPWVARLVFYPADRVMSVCLRVLYGCLWRDGLLSYAPV 180  
DB 140 EVISGNEDPEGVVLKDLGPPWVARLVFYPADRVMSVCLRVLYGCLWRDGLLSYAPV 199  
QY 181 GQTMYLSEA-YLNDSTYDGHVGLQGLQGLADGVGLDFFRSQELRWVPGYDYVG 239  
DB 200 GQTMQLSEVMVHLNDSTYDGHVGLQGLQGLADGVGLDFFRSQELRWVPGYDYVG 259  
QY 240 WSNISFSSGYVEMEFEDRLRAFAQMCHNMMHTLGLARLPGGVCECRFGPAMAWEGEP 299

Db 260 WSNQSEPTCYVEMEFEDRLRAFAQMCHNMMHTLGLARLPGGVCECRFGPAMAWEGEP 319  
QY 300 MHNHNGNGLDPRARVSVPLGGRVAFLOCRFLFAGPWLFLSEISFSDVNNSSPALG 359  
DB 320 VRHALGSLGSDPRARVSVPLGGRVAFLOCRFLFAGPWLFLSEISFSDVNNSSPALG 376  
QY 360 GTFPPAPWPPGPPPTNFSSLELEPRGQOPVAKPEGSPTA 399  
DB 376 DTFPPAPWPPGPPPTNFSSLELEPRGQOPVAKPEGSPTA 415

RESULT 4  
ID TRK3\_HUMAN STANDARD; PRT: 855 AA.  
AC Q16832;  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE RECEPTOR PROTEIN-TYROSINE KINASE TKT PRECURSOR (EC 2.7.1.112)  
DE (TYROSINE-PROTEIN KINASE TYRO 10) (NEUROTROPHIC TYROSINE KINASE, RECEPTOR-RELATED 3)  
DE NTKR3 OR TKT OR TYRO10.  
GN NTKR3 OR TKT OR TYRO10.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=HEART AND THYMUS;  
RX MEDLINE; 94067796.  
RA KARN T., HOLTRICH U., BRAEUNINGER A., BOEHME B., WOLF G., RUEBSAMEN-WAIGMANN H., STREBHARDT K.;  
RT "Structure, expression and chromosomal mapping of TKT from man and mouse: a new subclass of receptor tyrosine kinases with a factor VIII-like domain";  
RT ONCOGENE 8:3433-3440(1993).  
RL -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE -  
CC ADP + PROTEIN TYROSINE PHOSPHATE.  
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
CC -1- TISSUE SPECIFICITY: THE MAJOR 10 KD TRANSCRIPT IS EXPRESSED IN HIGH LEVELS IN HEART AND LUNG, LESS IN BRAIN, PLACENTA, LIVER, SKELETAL MUSCLE, PANCREAS, AND KIDNEY.  
CC -1- ALTERNATIVE PRODUCTS: DIFFERENT TRANSCRIPTS ARE DERIVED FROM ONE GENE.  
CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 F5/8 TYPE C DOMAIN.  
CC -1- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-PROTEIN KINASES.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: X74764; G433338; -  
CC MIN: 191311; -  
DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE: PS00239; RECEPTOR\_TYR\_KIN\_II; 1.  
DR PROSITE: PS01285; FA58C\_1; 1.  
DR PROSITE: PS01286; FA58C\_2; 1.  
DR PROSITE: PS00111; PROTEIN\_KINASE\_DOM; 1.  
DR PFAM: PF00069; pkinase; 1.  
DR PFAM: PF00754; F5\_F8\_type\_C; 1.  
DR HSP: P11362; 1FGI.  
KW TRANSFERASE: TYROSINE-PROTEIN KINASE; GLYCOPROTEIN; SIGNAL;  
KW PHOSPHORYLATION: TRANSMEMBRANE; RECEPTOR: ATP-BINDING.  
FT SIGNAL 1 21 POTENTIAL.  
FT CHAIN 22 855 RECEPTOR PROTEIN-TYROSINE KINASE TKT.  
FT DOMAIN 22 399 EXTRACELLULAR (POTENTIAL).  
FT DOMAIN 22 399

FT TRANSMEM 400 421 POTENTIAL.  
 FT DOMAIN 422 855 CYTOPLASMIC (POTENTIAL).  
 FT F5/8 TYPE C. 185  
 FT DOMAIN 30 185  
 FT NP\_BIND 569 577 PROTEIN KINASE.  
 FT BINDING 608 608 ATP (BY SIMILARITY).  
 FT ACT\_SITE 710 710 ATP (BY SIMILARITY).  
 FT DISULFID 30 185 BY SIMILARITY.  
 FT MOD\_RES 740 740 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CARBOHYD 121 121 POTENTIAL.  
 FT CARBOHYD 213 213 POTENTIAL.  
 FT CARBOHYD 261 261 POTENTIAL.  
 FT CARBOHYD 280 280 POTENTIAL.  
 FT CARBOHYD 372 372 POTENTIAL.  
 SQ SEQUENCE 855 AA; 96752 MW; 2623D841 CRC32;

Query Match 46.1%; Score 1002.5; DB 1; Length 855;  
 Best Local Similarity 53.6%; Pred. No. 6.4e-75;  
 Matches 196; Conservative 54; Mismatches 109; Indels 7; Gaps 6;

QY 5 KGHFPAKCRYALGMQDRTIPDSASSWSDSSTAARHSRLSSDGDGAWCPAGSVFPK 64  
 Db 22 KAQVNPACRYPLGSGGGIPDEDITASSOWSESTAAYGRDLSEGDGAWCPPIPVDP 81  
 QY 65 E-EYLVQDQLRLHLVALVGTQGRHAGGLGKEFSRSLRYSDRGRWGWKDRWQVEVI 123  
 Db 82 DLKEFLQIDLRLHLFTLVTQGRHAGGLGKEFSRSLRYSDRGRWGWKDRWQVEVI 141  
 QY 124 SGNPEGVVLKDLGPPMVARLVRYPRADRVMSCLVRLVGLCWRDGLLSYAPVQOT 183  
 Db 142 DGNPNPYDIFLADLEPPIVARVRIPTVDHSMNVCMLVGLCWRDGLLSYAPVQOT 201  
 QY 184 MYL--SEAVYLDNDYDHTVGLGQGLQGLADGVGLDFFRKSQELRVWPGDYDVGWS 241  
 Db 202 FVLPGGSIYLDNDYDHTVGLGQGLQGLADGVGLDFFRKSQELRVWPGDYDVGWR 260  
 QY 242 NSEFSGVYMEFEFDRLEAFQAMQVHNNHTLIGARLPGGVECFRRGRPMANWEGEPMR 301  
 Db 261 NESATNGIEINFEFDRINFTFMVHCNNMFAGVKIFKEVQCYF-RSEASEWEPNAIS 319  
 QY 302 HNGLGNLGDPRARVSVPLGGVRVFLQCRFLFAGPWLLEIFSEIFSD-VYNNSSPALGG 360  
 Db 320 FPLVLDNDVNPASFVTVPLHHRMSAIAKQYHFADTWMFSEITFQSDAAMINSEAL-P 378  
 QY 361 FPPAP 366  
 Db 379 TSPMAP 384

RESULT 5  
 TRK3\_MOUSE STANDARD; PRT; 854 AA.  
 ID TRK3\_MOUSE  
 AC Q62371;  
 DT 01-NOV-1997 (REL. 35, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
 DE RECEPTOR PROTEIN-TYROSINE KINASE TKT PRECURSOR (EC 2.7.1.112)  
 DE (TYROSINE-PROTEIN KINASE TYRO 10) (NEUROTROPHIC TYROSINE KINASE,  
 DE RECEPTOR-RELATED 3).  
 GN NTRK3 OR TKT OR TYRO10.  
 OS MUS MUSCULUS (MOUSE).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 94067796.  
 RA KARN T., HOLTRICH U., BRAEUNINGER A., BOEHME B., WOLF G.,  
 RA RUEBSAMEN-WALTMANN H., STREIBHARDT K.;  
 RT "Structure, expression and chromosomal mapping of TKT from man and  
 RT mouse: a new subclass of receptor tyrosine kinases with a factor  
 RT VIII-like domain."  
 RL ONCOGENE 8:3433-3440(1993).

RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-C57BL/6; TISSUE-BRAIN;  
 RX MEDLINE; 94151011.  
 RA LAI C., LEMKE G.E.;  
 RT "Structure and expression of the Tyro 10 receptor tyrosine kinase."  
 RL ONCOGENE 9:877-883(1994).  
 CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE -  
 ADP + PROTEIN TYROSINE PHOSPHATE.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED; HIGH LEVELS IN SKELETAL  
 MUSCLE, HEART, CNS, AND KIDNEY; LESS IN OTHER TISSUES. THE MAJOR  
 10 KD TRANSCRIPT IS EXPRESSED IN HIGH LEVELS IN HEART AND LUNG,  
 CC LESS IN BRAIN AND TESTIS.  
 CC -1- ALTERNATIVE PRODUCTS: DIFFERENT TRANSCRIPTS ARE DERIVED FROM ONE  
 GENE.  
 CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC  
 DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 F5/8 TYPE C DOMAIN.  
 CC -1- SIMILARITY: BELONGS TO THE INSULIN RECEPTOR FAMILY OF TYROSINE-  
 PROTEIN KINASES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X76505; G435162; ALT\_INIT.  
 DR MGD; MGI:97385; NTRK3.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS00239; RECEPTOR\_TYR\_KIN\_II; 1.  
 DR PROSITE; PS01285; FA38C\_1; 1.  
 DR PROSITE; PS01286; FA38C\_2; 1.  
 DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PFAM; PF00069; pkinase; 1.  
 DR PFAM; PF00754; F5\_F8\_type\_C; 1.  
 DR HSP; P11362; LFGI.  
 KW TRANSFERASE; TYROSINE-PROTEIN KINASE; GLYCOPROTEIN; SIGNAL;  
 KW PHOSPHORYLATION; TRANSMEMBRANE; RECEPTOR; ATP-BINDING.  
 FT SIGNAL 1 21  
 FT CHAIN 22 854 RECEPTOR PROTEIN-TYROSINE KINASE TKT.  
 FT DOMAIN 22 399 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 400 421 POTENTIAL.  
 FT DOMAIN 422 854 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 30 185 F5/8 TYPE C.  
 FT DOMAIN 563 ? PROTEIN KINASE.  
 FT NP\_BIND 569 577 ATP (BY SIMILARITY).  
 FT BINDING 608 608 ATP (BY SIMILARITY).  
 FT ACT\_SITE 709 709 BY SIMILARITY.  
 FT DISULFID 30 185 BY SIMILARITY.  
 FT MOD\_RES 739 739 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT CARBOHYD 121 121 POTENTIAL.  
 FT CARBOHYD 213 213 POTENTIAL.  
 FT CARBOHYD 261 261 POTENTIAL.  
 FT CARBOHYD 280 280 POTENTIAL.  
 FT CARBOHYD 372 372 POTENTIAL.  
 SQ SEQUENCE 854 AA; 96482 MW; B78B6551 CRC32;

Query Match 45.8%; Score 995.5; DB 1; Length 854;  
 Best Local Similarity 49.6%; Pred. No. 2.4e-74;  
 Matches 198; Conservative 58; Mismatches 116; Indels 27; Gaps 7;  
 QY 5 KGHFPAKCRYALGMQDRTIPDSASSWSDSSTAARHSRLSSDGDGAWCPAGSVFPK 64  
 Db 22 KAQVNPACRYPLGSGGGIPDEDITASSOWSESTAAYGRDLSEGDGAWCPPIPVDP 81  
 QY 65 E-EYLVQDQLRLHLVALVGTQGRHAGGLGKEFSRSLRYSDRGRWGWKDRWQVEVI 123  
 Db 82 DLKEFLQIDLRLHLFTLVTQGRHAGGLGKEFSRSLRYSDRGRWGWKDRWQVEVI 141





```
FT DOMAIN 27 141 CUB 1.
FT DOMAIN 147 265 CUB 2.
FT DOMAIN 275 424 F5/8 TYPE C 1.
FT DOMAIN 431 583 F5/8 TYPE C 2.
FT DOMAIN 645 811 MAM.
FT DISULFID 27 54 PROBABLE.
FT DISULFID 82 104 PROBABLE.
FT DISULFID 147 173 PROBABLE.
FT DISULFID 206 228 PROBABLE.
FT DISULFID 275 424 BY SIMILARITY.
FT DISULFID 431 583 BY SIMILARITY.
FT CARBOHYD 150 150 POTENTIAL.
FT CARBOHYD 261 261 POTENTIAL.
FT CARBOHYD 300 300 POTENTIAL.
FT CARBOHYD 522 522 POTENTIAL.
FT CARBOHYD 842 842 POTENTIAL.
SQ SEQUENCE 923 AA; 103019 MW; 265CODE7 CRC32;

Query Match 9.4%; Score 204; DB 1; Length 923;
Best Local Similarity 31.1%; Pred. No. 3.4e-09;
Matches 56; Conservative 26; Mismatches 72; Indels 26; Gaps 6;

QY 12 KCRYALGMODRTIPDSISASSWSDSSTAARHSRLSSDGDGAWCPAGSVFPKKEEYLVQV 71
DB 274 KCMEALGMESGEIHSQDITASSQYGTNNVSRLNYP--ENGWTPGDSY---KEWTVQ 328
QY 72 DLQRLHLVALVGTQGHAGGLCKE-----FSRSYRLYSRQGRRWGKWRGQEVISGN 126
DB 329 DLGLLRFVTAVGTC---GAISKETKKYKTVYKVDISSNGEDWISLKEGNKALIFQGN 384
QY 127 EDPEGVVLKDLGPPMVARLVRYPRADRVMSVCLRVLYGCLWRD-----GLLS 175
DB 385 TNPDTVLGVFSEKPLITRFVRIKP-VSNETGISMRFEVYGCKITDPCSGMLGMVSGILS 443

RESULT 9
ID F88_HUMAN STANDARD; PRT: 2351 AA.
AC P00451.
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)
DE COAGULATION FACTOR VIII PRECURSOR (PROCOAGULANT COMPONENT).
GN F8C.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 86081164.
RA TRUETT M.A., BLACHER R., BURKE R.L., CAPUT D., CHU C., DINA D.,
RA HARTOG K., KUO C.H., MASARZ F.R., MERRYWEATHER J.P., NAJARIAN R.,
RA PACHL C., POTTER S.J., POMA J., QUIROGA M., RALL L.B., RANDOLPH A.,
RA URDEA M.S., VALENZUELA P., DAHL H.-H.M., FAVALLARO J., HANSEN J.,
RA NORDFANG O., EZBAN M.;
RA "Characterization of the polypeptide composition of human factor
RT VIII:C and the nucleotide sequence and expression of the human kidney
RT cDNA."
RL DNA 4:333-349(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 85061548.
RA WOOD W.I., CAPON D.J., SIMONSEN C.C., EATON D.L., GITSCHIER J.,
RA KEYS B., SEEBURG P.H., SMITH D.H., HOLLINGSHEAD P., WION K.L.,
RA DELMART E., TUDDENHAM E.G.D., VEHAR G.A., LAWN R.M.;
RA "Expression of active human factor VIII from recombinant DNA clones."
RL NATURE 312:330-337(1984).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE: 85061550.
RA TOOLE J.J., KNOPP J.L., WOZNEY J.M., SULTZMAN L.A., BUECKER J.L.,
RA PITTMAN D.D., KAUFMAN R.J., BROWN E., SHOEMAKER C., ORR E.C.,
```

```
RA AMPHLETT G.W., FOSTER W.B., COE M.L., KNUTSON G.J., PASS D.N.,
RA HEWICK R.M.;
RT "Molecular cloning of a cDNA encoding human antihemophilic factor."
RL NATURE 312:342-347(1984).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE: 93265012.
RA GITSCHIER J., WOOD W.I.;
RT "Sequence of the exon-containing regions of the human factor VIII
RT gene."
RL HUM. MOL. GENET. 1:199-200(1992).
RN [5]
RP SEQUENCE OF 2064-2070 FROM N.A.
RA DE WATER N.S., WILLIAMS R., BROWETT P.J.;
RL SUBMITTED (JUN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [6]
RP SULFATATION OF 1699.
RX MEDLINE: 91093266.
RA LEYTE A., VAN SCHIJNDEL H.B., NIEHRS C., HUTTNER W.B., VERBEET M.P.,
RA MERTENS K., VAN MOURIK J.A.;
RT "Sulfation of Tyr1680 of human blood coagulation factor VIII is
RT essential for the interaction of factor VIII with von Willebrand
RT factor."
RL J. BIOL. CHEM. 266:740-746(1991).
RN [7]
RP SULFATATION.
RX MEDLINE: 92207952.
RA PITTMAN D.D., WANG J.H., KAUFMAN R.J.;
RT "Identification and functional importance of tyrosine sulfate
RT residues within recombinant factor VIII."
RL BIOCHEMISTRY 31:3315-3325(1992).
RN [8]
RP STRUCTURE BY NMR OF 2322-2343.
RX MEDLINE: 95200924.
RA GILBERT G.E., BALEJA J.D.;
RT "Membrane-binding peptide from the C2 domain of factor VIII forms an
RT amphipathic structure as determined by NMR spectroscopy."
RL BIOCHEMISTRY 34:3022-3031(1995).
RN [9]
RP REVIEW ON MOLECULAR BASIS OF HEMOPHILIA A.
RX MEDLINE: 91221499.
RA GITSCHIER J.;
RT "The molecular basis of hemophilia A."
RL ANN. N.Y. ACAD. SCI. 614:89-96(1991).
RN [10]
RP REVIEW ON MOLECULAR BASIS OF HEMOPHILIA A.
RX MEDLINE: 89088506.
RA WHITE G.C. II, SHOEMAKER C.B.;
RT "Factor VIII gene and hemophilia A."
RL BLOOD 73:11-12(1989).
RN [11]
RP REVIEW ON MOLECULAR BASIS OF HEMOPHILIA A.
RX MEDLINE: 95245332.
RA ANTONARAKIS S.E., KAZAZIAN H.H., TUDDENHAM E.G.D.;
RT "Molecular etiology of factor VIII deficiency in hemophilia A."
RL HUM. MUTAT. 5:1-22(1995).
RN [12]
RP VARIANT GLN-2326.
RX MEDLINE: 86235434.
RA GITSCHIER J., WOOD W.I., SHUMAN M.A., LAWN R.M.;
RT "Identification of a missense mutation in the factor VIII gene of a
RT mild hemophilic."
RL SCIENCE 232:1415-1416(1986).
RN [13]
RP VARIANT PRO-2135.
RX MEDLINE: 88096539.
RA LEVINSON B., JANCO R.L., PHILLIPS J.A. III, GITSCHIER J.;
RT "A novel missense mutation in the factor VIII gene identified by
RT analysis of amplified hemophilia DNA sequences."
RL NUCLEIC ACIDS RES. 15:9797-9805(1987).
RN [14]
RP VARIANT GLN-2228.
RX MEDLINE: 88191889.
```



modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; D45416; G1841311; --  
DR PROSITE; PS01180; CUB; 2.  
DR PROSITE; PS01285; FA58C.1; 2.  
DR PROSITE; PS01286; FA58C.2; 2.  
DR PROSITE; PS00740; MAM.1; 1.  
DR PROSITE; PS50060; MAM.2; 1.  
DR PFAM; PF00431; CUB; 2.  
DR PFAM; PF00629; MAM; 1.  
DR PFAM; PF00754; F5F8\_type\_C; 2.  
KW TRANSMEMBRANE; GLYCOPROTEIN; NEURONE; SIGNAL; REPEAT; CELL ADHESION.  
FT SIGNAL 1 18  
FT CHAIN 19 914  
FT DOMAIN 20 847  
FT DOMAIN 848 870  
FT DOMAIN 871 914  
FT DOMAIN 25 139  
FT DOMAIN 145 263  
FT DOMAIN 273 422  
FT DOMAIN 429 581  
FT DOMAIN 636 801  
FT DOMAIN 852 914  
FT DISULFID 25 52  
FT DISULFID 80 102  
FT DISULFID 145 171  
FT DISULFID 204 226  
FT DISULFID 273 422  
FT DISULFID 429 581  
FT SEQUENCE 914 AA; 102480 MW; 593B7139 CRC32;

Query Match 8.7%; Score 189; DB 1; Length 914;

Best Local Similarity 31.3%; Pred No. 5,7e-08;  
Matches 56; Conservative 26; Mismatches 83; Indels 14; Gaps 5;

QY 12 KCRYALGMQDRTIPDSISASSWSDSTAARSHRLESSDGDGAWCPAGSVFPKKEEYQV 71  
DB 272 QCMPLGMEGSHDSQITVSQYSAINSSERSRLNYP--ENGWTPGED---SVREWIQV 326  
QY 72 DLQRLHLVALVGTQGRHAGLCKE-FRSYRLRYSDGRWGWKDRGQVEISGNEDE 130  
DB 327 DLGLRFLVSGIGTQGAISKETKYLKTYRVDVSSNGEDWTLKGNKPVVFGQNSPT 386  
QY 131 GYVLKDLGPPVAVRLRYFPRADRVMSVCLRVLYGCLWRD-----GLLSYAPVQV 182  
DB 387 DVYRFPKPLTRFRIKP-VSWENGVSURFEVYGCKITDYPGSMGLGWGLIPDSQ 444

RESULT 11

ID FA5\_HUMAN STANDARD; PRT; 2224 AA.  
AC P12259; Q14285;  
DT 01-OCT-1989 (REL. 12, CREATED)  
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)  
DE COAGULATION FACTOR V PRECURSOR (ACTIVATED PROTEIN C COFACTOR).  
GN F5.  
OS HOMO SAPIENS (HUMAN).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 92232668.  
RA CRIFE L.D., MOORE K.D., KANE W.H.;  
RT "Structure of the gene for human coagulation factor V.";  
RL BIOCHEMISTRY 31:3777-3785(1992).  
[2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 87260886.  
RA JENNY R.J., PITTMAN D.D., TOOLE J.J., KRIZ R.W., ALDAPE R.A.,  
RA HEWICK R.M., KAUFMAN R.J., MANN K.G.;

RT "Complete cDNA and derived amino acid sequence of human factor V.";  
RL PROC. NATL. ACAD. SCI. U.S.A. 84:4846-4850(1987).  
RN [3]  
RX SEQUENCE OF 1-1600 FROM N.A.  
RX MEDLINE; 88107560.  
RA KANE W.H., ICHINOSE A., HAGEN F.S., DAVIE E.W.;  
RT "Cloning of cDNAs coding for the heavy chain region and connecting  
RT region of human factor V, a blood coagulation factor with four types  
RT of internal repeats.";  
RL BIOCHEMISTRY 26:6508-6514(1987).  
RN [4]  
RP SEQUENCE OF 1188-1215 AND 1315-2224 FROM N.A.  
RX MEDLINE; 86313665.  
RA KANE W.H., DAVIE E.W.;  
RT "Cloning of a cDNA coding for human factor V, a blood coagulation  
RT factor homologous to factor VIII and ceruloplasmin.";  
RL PROC. NATL. ACAD. SCI. U.S.A. 83:6800-6804(1986).  
RN [5]  
RP PARTIAL SEQUENCE FROM N.A.  
RX TISSUE-FIBROBLAST;  
RX MEDLINE; 93203619.  
RA SHEN N.L., FAN S.-T., PYATI J., GRAFF R., LAPOLLA R.J.,  
RT "The serine protease cofactor factor V is synthesized by  
RT lymphocytes.";  
RL J. IMMUNOL. 150:2992-3001(1993).  
RN [6]  
RP VARIANT APCR GIN-534.  
RX MEDLINE; 94217810.  
RA BERTINA R.M., KOELEMAN B.P.C., KOSTER T., ROSENDAAL F.R.,  
RA DIRVEN R.J., DE RONDE H., VAN DER VELDEN P.A., REITSMA P.H.;  
RT "Mutation in blood coagulation factor V associated with resistance to  
RT activated protein C.";  
RL NATURE 369:64-67(1994).  
RN [7]  
RP FUNCTION: COAGULATION FACTOR V IS A COFACTOR THAT PARTICIPATES  
CC WITH FACTOR XA TO ACTIVATE PROTHROMBIN TO THROMBIN.  
CC -I- SUBUNIT: FACTOR VA IS COMPOSED OF AN HEAVY CHAIN AND OF A LIGHT  
CC CHAIN NONCOVALENTLY BOUND. THE INTERACTION BETWEEN THE TWO CHAINS  
CC IS CALCIUM-DEPENDENT.  
CC -I- DOMAIN: DOMAIN B CONTAINS 35 X 9 AA TANDEM REPEATS, AND 2 X 17 AA  
CC REPEATS.  
CC -I- PTM: THROMBIN ACTIVATES FACTOR V PROTEOLYTICALLY TO THE ACTIVE  
CC COFACTOR, FACTOR V(A) (FORMATION OF A HEAVY CHAIN AT THE N-  
CC TERMINUS AND A LIGHT CHAIN AT THE C-TERMINUS).  
CC -I- DISEASE: OMEN PARAHEMOPHILIA, AN HEMORRHAGIC DIATHESIS, IS DUE  
CC TO A DEFICIENCY OF FACTOR V. OTHER DEFECTS IN F5 RESULTS IN A  
CC FORM OF THROMBOPHILIA KNOWN AS APC RESISTANCE (APCR). THE APCR  
CC MUTATION IS FOUND IN ABOUT 5% OF THE POPULATION WHICH SUGGEST THAT  
CC A SLIGHT THROMBOTIC TENDENCY MAY CONFER SOME ADVANTAGE IN FETAL  
CC IMPLANTATION.  
CC -I- SIMILARITY: CONTAINS 3 F5/8 TYPE A DOMAINS; EACH IS COMPOSED OF  
CC 2 PLASTOCYANIN-LIKE REPEATS.  
CC -I- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.  
CC -I- SIMILARITY: STRONG, TO COAGULATION FACTOR VIII.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; L32779; G488110; JOINED.  
DR EMBL; L32755; G488110; JOINED.  
DR EMBL; L32756; G488110; JOINED.  
DR EMBL; L32757; G488110; JOINED.  
DR EMBL; L32758; G488110; JOINED.  
DR EMBL; L32759; G488110; JOINED.  
DR EMBL; L32760; G488110; JOINED.  
DR EMBL; L32761; G488110; JOINED.  
DR EMBL; L32762; G488110; JOINED.  
DR EMBL; L32763; G488110; JOINED.



DR EMBL; L32764; G488110; JOINED.  
 DR EMBL; L32765; G488110; JOINED.  
 DR EMBL; L32766; G488110; JOINED.  
 DR EMBL; L32767; G488110; JOINED.  
 DR EMBL; L32768; G488110; JOINED.  
 DR EMBL; L32769; G488110; JOINED.  
 DR EMBL; L32770; G488110; JOINED.  
 DR EMBL; L32771; G488110; JOINED.  
 DR EMBL; L32772; G488110; JOINED.  
 DR EMBL; L32773; G488110; JOINED.  
 DR EMBL; L32774; G488110; JOINED.  
 DR EMBL; L32775; G488110; JOINED.  
 DR EMBL; L32776; G488110; JOINED.  
 DR EMBL; L32777; G488110; JOINED.  
 DR EMBL; L32778; G488110; JOINED.  
 DR EMBL; M16967; G182412; -.  
 DR EMBL; M14335; G182798; -.  
 DR PIR; A25897; A25897.  
 DR PIR; A28028; A28028.  
 DR MIM; 134400; -.  
 DR MIM; 188055; -.  
 DR MIM; 227310; -.  
 DR MIM; 227400; -.  
 DR PROSITE; PS00079; MULTICOPPER\_OXIDASE1; 2.  
 DR PROSITE; PS01285; FA58C\_1; 2.  
 DR PROSITE; PS01286; FA58C\_2; 2.  
 DR PFAM; PF00394; Cu-oxidase; 3.  
 DR PFAM; PF00754; F5\_F8\_type\_C; 2.  
 DR HSP; P00450; 1KCW.  
 KW BLOOD COAGULATION; PLASMA; GLYCOPROTEIN; CALCIUM; SIGNAL; ZYMOGEN;  
 REPEAT; POLYMORPHISM; DISEASE MUTATION; THROMBOPHILIA.  
 FT SIGNAL 1 28  
 FT CHAIN 29 2224  
 FT PEPTIDE 29 737  
 FT CHAIN 1738 1573  
 FT CHAIN 1574 2224  
 FT DOMAIN 30 329  
 FT DOMAIN 30 193  
 FT DOMAIN 203 329  
 FT DOMAIN 348 684  
 FT DOMAIN 348 526  
 FT DOMAIN 536 684  
 FT DOMAIN 692 1573  
 FT DOMAIN 895 928  
 FT DOMAIN 895 911  
 FT REPEAT 912 928  
 FT REPEAT 1135 1148  
 FT SIMILAR 1185 1463  
 FT DOMAIN 1185 1463  
 FT REPEAT 1185 1193  
 FT REPEAT 1194 1202  
 FT REPEAT 1203 1211  
 FT REPEAT 1212 1220  
 FT REPEAT 1221 1229  
 FT REPEAT 1230 1238  
 FT REPEAT 1239 1247  
 FT REPEAT 1248 1256  
 FT REPEAT 1257 1265  
 FT REPEAT 1266 1274  
 FT REPEAT 1275 1283  
 FT REPEAT 1284 1292  
 FT REPEAT 1293 1301  
 FT REPEAT 1302 1310  
 FT REPEAT 1311 1319  
 FT REPEAT 1320 1328  
 FT REPEAT 1329 1337  
 FT REPEAT 1338 1346  
 FT REPEAT 1347 1355  
 FT REPEAT 1356 1364  
 FT REPEAT 1365 1373  
 FT REPEAT 1374 1382  
 FT REPEAT 1383 1391  
 FT REPEAT 1392 1400

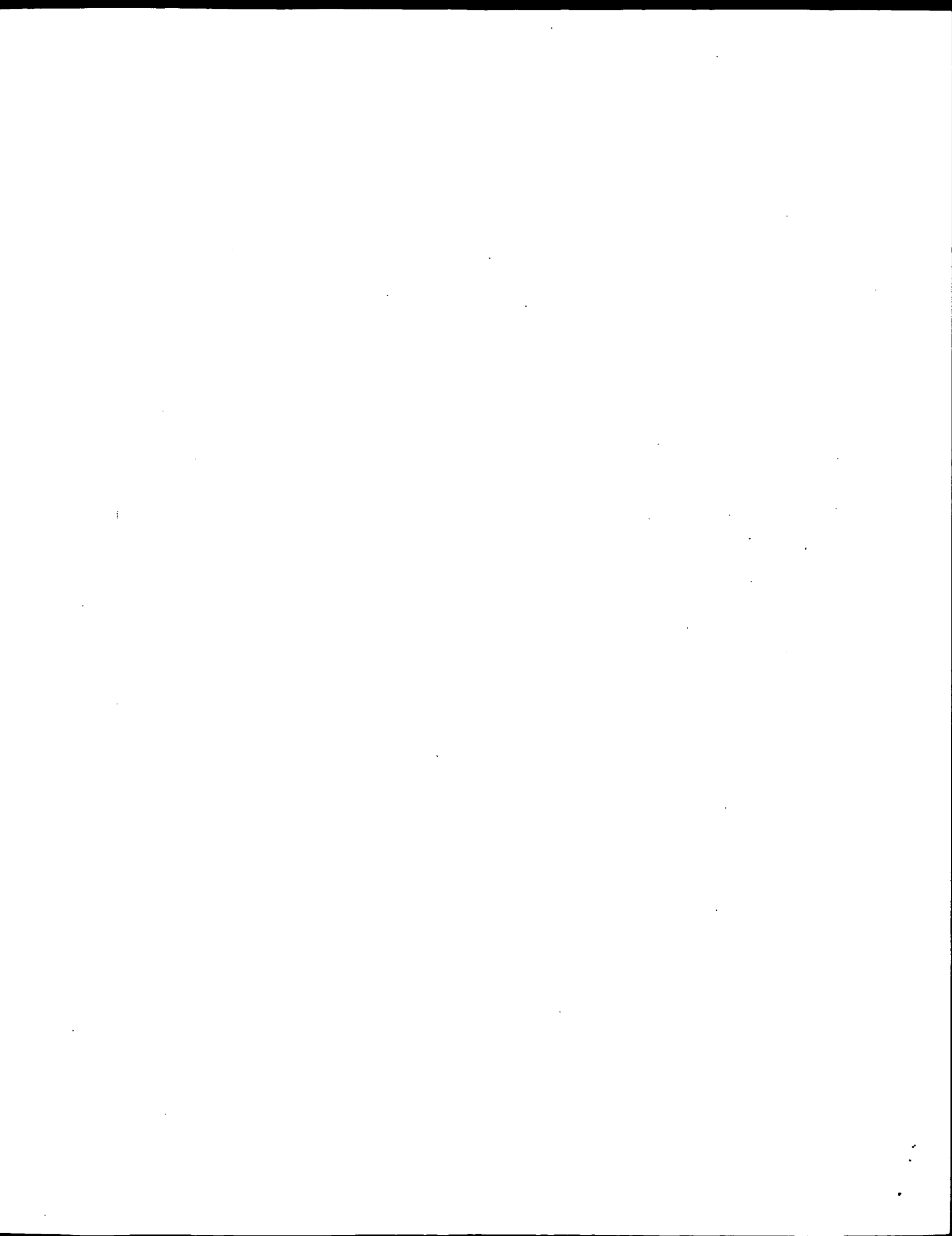
FT REPEAT 1401 1409 25.  
 FT REPEAT 1410 1418 26.  
 FT REPEAT 1419 1427 27.  
 FT REPEAT 1428 1436 28.  
 FT REPEAT 1437 1445 29.  
 FT REPEAT 1446 1454 30.  
 FT REPEAT 1455 1463 31.  
 FT REPEAT 1464 1472 32.  
 FT REPEAT 1473 1481 33.  
 FT REPEAT 1482 1490 34.  
 FT REPEAT 1493 1501 35.  
 FT DOMAIN 1578 1907 F5/8 TYPE A 3.  
 FT DOMAIN 1751 1907 PLASTOCYANIN-LIKE 5.  
 FT DOMAIN 1761 1907 PLASTOCYANIN-LIKE 6.  
 FT DOMAIN 1907 2061 F5/8 TYPE C 1.  
 FT DOMAIN 2066 2221 F5/8 TYPE C 2.  
 FT SITE 737 738 CLEAVAGE (BY THROMBIN).  
 FT SITE 1046 1047 CLEAVAGE (BY THROMBIN).  
 FT SITE 1573 1574 CLEAVAGE (BY THROMBIN).  
 FT DISULFID 167 193 PROBABLE.  
 FT DISULFID 500 526 PROBABLE.  
 FT DISULFID 1725 1751 PROBABLE.  
 FT DISULFID 1907 2061 BY SIMILARITY.  
 FT DISULFID 2066 2221 BY SIMILARITY.  
 FT CARBOHYD 51 51 POTENTIAL.  
 FT CARBOHYD 55 55 POTENTIAL.  
 FT CARBOHYD 239 239 POTENTIAL.  
 FT CARBOHYD 297 297 POTENTIAL.  
 FT CARBOHYD 382 382 POTENTIAL.  
 FT CARBOHYD 460 460 POTENTIAL.  
 FT CARBOHYD 468 468 POTENTIAL.  
 FT CARBOHYD 554 554 POTENTIAL.  
 FT CARBOHYD 741 741 POTENTIAL.  
 Query Match 8.2%; Score 178.5; DB 1; Length 2224;  
 Best Local Similarity 28.4%; Pred. No. 1.2e-06;  
 Matches 46; Conservative 32; Mismatches 71; Indels 13; Gaps 6;  
 QY 13 CRYALGMQDRTIPDSISASSS---WSDTAARHSRLESSDDGAWCPAGSVFPKEEY 68  
 Db 2066 CSTPLGMEKNGKIENKQITASSPKKSWMGDYWEPFRARNAQGRVNAWAKAN---NNKQW 2122  
 QY 69 LQVLDLQRLHLVALVGTQGRHAGLGKE-FSRSYRLYSRDGRRWG--KDRWGQEVISG 125  
 Db 2123 LEIDLKIKKITAIIQ--CKSLSEMVKSVYTHYSQGVKWPYRLKSSNVDKIFEG 2180  
 QY 126 NEDPEGVVLKDLGPPMVARLVFRYPADRVMSVCLVELYGC 167  
 Db 2181 NTNTKGHVKNFNPPIISRFIRVPIKTN-QSITLRLELFGC 2221  
 RESULT 12  
 MFGM HUMAN  
 ID MFGM\_HUMAN STANDARD; PRT; 387 AA.  
 AC Q08431;  
 DT 01-OCT-1996 (REL. 34, CREATED)  
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)  
 DE MILK FAT GLOBULE-EGF FACTOR 8 PRECURSOR (MFG-E8) (HMFG) (BREAST  
 DE EPITHELIAL ANTIGEN BA46) (MFGM).  
 GN MFG8.  
 OS HOMO SAPIENS (HUMAN).  
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BREAST, AND BREAST CARCINOMA;  
 RX MEDLINE; 96213908.  
 RA COUTO J.R., TAYLOR M.R., GODWIN S.G., CERIANI R.L., PETERSON J.A.;  
 RT "Cloning and sequence analysis of human breast epithelial antigen  
 RT BA46 reveals an RGD cell adhesion sequence presented on an epidermal  
 RT growth factor-like domain";  
 RT DNA CELL BIOL. 15:281-286(1996).





RESULT	15
MFGM_MOUSE	
ID	MFGM_MOUSE
AC	P21956
DT	01-AUG
DT	01-AUG
DT	01-NOV
DE	MIYAKI





GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 4, 1999, 01:52:29 ; Search time 19.79 Seconds  
(without alignments)  
807.788 Million cell updates/sec

Title: US-08-170-558-8  
Perfect score: 2175  
Sequence: 1 DADMGHFDPAKCRYALGMQ.....LELEPRGQOPVAKPEGSPTA 399

Scoring table: BLOSUM62

Searched: 122810 seqs, 40065486 residues

Database: PIR\_60.\*

1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2167	99.6	876	2 A49508	protein-tyrosine k
2	2164	99.5	913	2 A48280	receptor tyrosine
3	2021	92.9	910	2 A53137	tyrosine kinase re
4	1002.5	46.1	855	2 S42621	protein-tyrosine k
5	952.5	43.8	819	2 I48859	tyro 10 receptor k
6	213	9.8	927	2 JQ0948	A5 antigen precurs
7	211.5	9.7	845	3 JC5256	adipocyte transcri
8	192	8.8	2351	1 E2HU	coagulation factor
9	182	8.8	216	2 A4258	factor VIII-associ
10	179.5	8.3	2224	1 KF0U5	coagulation factor
11	177.5	8.2	218	2 A47285	milk fat globule p
12	175	8.0	2319	2 A47004	coagulation factor
13	174.5	8.0	2211	1 KF905	coagulation factor
14	174.5	8.0	463	2 A36479	coagulation factor
15	172.5	7.9	427	2 JC4915	milk fat globule m
16	167.5	7.7	401	2 S65138	ags protein precu
17	167.5	7.7	427	2 S74211	glycoprotein antig
18	164.5	7.6	719	2 S51739	PAS-6/7 protein pr
19	145.5	6.7	3133	2 S52093	transcription repr
20	106.5	4.9	448	2 A57435	hemocytin - silkwo
21	97	4.5	253	1 DLBOIA	3,4-dihydroxypheny
22	97	4.5	1269	2 S35366	discoicidin I chain
23	97	4.5	899	2 S17546	furl protein - fru
24	96	4.4	253	1 DLBOIC	probable serine pr
25	94.5	4.3	395	2 S76793	discoicidin I chain
26	92	4.2	635	2 S74718	hypothetical prote
27	89	4.1	3164	1 WMBEH6	sulfite reductase
28	89	4.1	477	2 S2162	UL36 protein - hum
29	88	4.0	1384	3 T02748	sucrose hydrolase
30	87.5	4.0	139	1 C45051	hypothetical prote
31	87.5	4.0	1168	1 MWAXIC	lamprin 2 precuso
32	87.5	4.0	420	2 A70796	myosin heavy chain
33	87	4.0	439	2 I60916	hypothetical prote
34	87	4.0	404	2 S66533	HNF-3/forhead hom
35	87	4.0	409	2 E69046	flavoprotein A, 45
36	86.5	4.0	149	1 DLBOID	flavoprotein AI -
37	86.5	4.0	377	2 A45851	discoicidin I chain
38	86.5	4.0	149	2 A03382	MHC class I histoc
39	86.5	4.0	379	2 S09791	discoicidin I chain
					hypothetical prote

ribosomal protein  
probable regulator  
DNA-directed RNA p  
hypothetical prote  
bicaudal-C - fruit  
hypothetical prote

#### ALIGNMENTS

##### RESULT 1

A49508  
protein-tyrosine kinase (EC 2.7.1.112) trke precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 10-Nov-1995 #sequence.revision 10-Nov-1995 #text\_change 16-Dec-1998  
C:Accession: A49508; I38358; S37402  
R:Di Marco, E.; Cutuli, N.; Guerra, L.; Cancedda, R.; De Luca, M.  
J. Biol. Chem. 268, 24290-24295, 1993  
A:Title: Molecular cloning of trke, a novel trk-related putative tyrosine kinase rece  
A:Reference number: A49508; MUID:94043265  
A:Accession: A49508  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-876 <DIA>  
A:Cross-references: EMBL:X74979; NID:g400462; PID:g400463  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; discoidin I amino-  
C:Keywords: Anp; phosphotransferase  
F:30-185/Domain: discoidin I amino-terminal homology <DNI>  
F:571-875/Domain: protein kinase homology <KIN>  
F:579-587/Region: protein kinase ATP-binding motif

Query Match 99.6% Score 2167; DB 2; Length 876;  
Best Local Similarity 99.7%; Pred. No. 2.2e-177;  
Matches 398; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	DADMGHFDPAKCRYALGMODRTIPDSISASSWSWSDTAARHSRLSSDGDGAWCPAGS	60
Db	19	DADMGHFDPAKCRYALGMODRTIPDSISASSWSWSDTAARHSRLSSDGDGAWCPAGS	78
Qy	61	VPKEEYLVQDLQRLHLVALVGTQGRHAGLGKEFSSRYRLYSRDRGRMWGKDRWGQ	120
Db	79	VPKEEYLVQDLQRLHLVALVGTQGRHAGLGKEFSSRYRLYSRDRGRMWGKDRWGQ	138
Qy	121	EVISGNEDEPGVWLKDGPPMVARLVRFYPRADRVMSVCLRVLYGCLWRDGLLSYAPV	180
Db	139	EVISGNEDEPGVWLKDGPPMVARLVRFYPRADRVMSVCLRVLYGCLWRDGLLSYAPV	198
Qy	181	GOTMYLSEAVYLVNDSTYDGHVTGGVGLQGLQGLADGVVGLDDFRKSOELRWMPGYDYVGW	240
Db	199	GOTMYLSEAVYLVNDSTYDGHVTGGVGLQGLQGLADGVVGLDDFRKSOELRWMPGYDYVGW	258
Qy	241	SNHFSFGYVMEFEFDRLRAFAQMVHCHNMHTLGLARLPGGVECFRRFRGPAMAWEGEPM	300
Db	259	SNHFSFGYVMEFEFDRLRAFAQMVHCHNMHTLGLARLPGGVECFRRFRGPAMAWEGEPM	318
Qy	301	RHNLGNNLGDPRARVSVPLGGVRARELQCRFLFAGPWLFFSEISFISDVVNNSSPALGG	360
Db	319	RHNLGNNLGDPRARVSVPLGGVRARELQCRFLFAGPWLFFSEISFISDVVNNSSPALGG	378
Qy	361	TFPPAPWMPGPPPTNFSSLELEPRGQOPVAKPEGSPTA	399
Db	379	TFPPAPWMPGPPPTNFSSLELEPRGQOPVAKPEGSPTA	417

##### RESULT 2

A48280  
receptor tyrosine kinase - human  
C:Species: Homo sapiens (man)  
C:Date: 31-May-1996 #sequence.revision 31-May-1996 #text\_change 16-Dec-1998  
C:Accession: A48280  
R:Johnson, J.D.; Edman, J.C.; Rutter, W.J.

Proc. Natl. Acad. Sci. U.S.A. 90, 5677-5681, 1993  
A:Title: A receptor tyrosine kinase found in breast carcinoma cells has an extracellular  
A:Reference number: A48280; MUID:93296201  
A:Accession: A48280  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-913 <RES>  
A:Cross-references: GB:L11315; NID:g403386; PID:g403387  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; discoidin I amino-ter  
F:30-185/Domain: discoidin I amino-terminal homology <DN1>  
F:608-912/Domain: protein kinase homology <KIN>  
F:616-624/Region: protein kinase ATP-binding motif

Query Match 99.5%; Score 2164; DB 2; Length 913;  
Best Local Similarity 99.5%; Pred. No. 4.2e-177;  
Matches 397; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 DADMGHFDPAKCRYALGMDRTIPDSISASSWSSTAARHSRLSSDGDGAWCPAGS 60  
DB 19 DADMGHFDPAKCRYALGMDRTIPDSISASSWSSTAARHSRLSSDGDGAWCPAGS 78  
QY 61 VFPKEEYLVQDLQRLHLVALVGTQGRHAGGLGKEFSRSLRYSRDRGRMMGWKDRWGQ 120  
DB 79 VFPKEEYLVQDLQRLHLVALVGTQGRHAGGLGKEFSRSLRYSRDRGRMMGWKDRWGQ 138  
QY 121 EVISGNEDEGPGVVLKDLGPPMVARLVREYPRADRVMSVCLRVLYGCLWRDGLLSYAPV 180  
DB 139 EVISGNEDEGPGVVLKDLGPPMVARLVREYPRADRVMSVCLRVLYGCLWRDGLLSYAPV 198  
QY 181 GOTMYLSEAVYLNDSYDGHVVGGLQGLADGVVGLDFFRKSQELRWMPGYDYVGW 240  
DB 199 GOTMYLSEAVYLNDSYDGHVVGGLQGLADGVVGLDFFRKSQELRWMPGYDYVGW 258  
QY 241 SNHSSFGYVEMEFDFRLRAFOAMVHCNNMHTLGARLPGGVECFRRGPPAMAWEGEPM 300  
DB 259 SNHSSFGYVEMEFDFRLRAFOAMVHCNNMHTLGARLPGGVECFRRGPPAMAWEGEPM 318  
QY 301 RHNLGNLGDPRARAVSVPLGGRVAREFQCRFLFAGPWLLESEIFSDVYNNSSPALGG 360  
DB 319 RHNLGNLGDPRARAVSVPLGGRVAREFQCRFLFAGPWLLESEIFSDVYNNSSPALGG 378  
QY 361 TFPAPMPPGPPPTNFSSLELEPRGQOPVAKPEGSPTA 399  
DB 379 TFPAPMPPGPPPTNFSSLELEPRGQOPVAKPEGSPTA 417

RESULT 3  
A53137  
tyrosine kinase receptor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 16-Dec-1998  
C:Accession: A53137  
R:Sanchez, M.P.; Tapley, P.; Saini, S.S.; He, B.; Pulido, D.; Barbacid, M.  
Proc. Natl. Acad. Sci. U.S.A. 91, 1819-1823, 1994  
A:Title: Multiple tyrosine protein kinases in rat hippocampal neurons: isolation of ptk-  
A:Reference number: A53137; MUID:94173920  
A:Accession: A53137  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-910 <RES>  
A:Cross-references: GB:L26525; NID:g432480; PID:g432481  
C:Genetics:  
A:Gene: ptk-3  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; discoidin I amino-ter  
F:31-186/Domain: discoidin I amino-terminal homology <DN1>  
F:605-909/Domain: protein kinase homology <KIN>  
F:613-621/Region: protein kinase ATP-binding motif

Query Match 92.9%; Score 2021; DB 2; Length 910;  
Best Local Similarity 93.0%; Pred. No. 7e-165;  
Matches 371; Conservative 9; Mismatches 15; Indels 4; Gaps 1;

QY 1 DADMGHFDPAKCRYALGMDRTIPDSISASSWSSTAARHSRLSSDGDGAWCPAGS 60  
DB 20 DADMGHFDPAKCRYALGMDRTIPDSISASSWSSTAARHSRLSSDGDGAWCPAGP 79  
QY 61 VFPKEEYLVQDLQRLHLVALVGTQGRHAGGLGKEFSRSLRYSRDRGRMMGWKDRWGQ 120  
DB 80 VFPKEEYLVQDLQRLHLVALVGTQGRHAGGLGKEFSRSLRYSRDRGRMMGWKDRWGQ 139  
QY 121 EVISGNEDEGPGVVLKDLGPPMVARLVREYPRADRVMSVCLRVLYGCLWRDGLLSYAPV 180  
DB 140 EVISGNEDEGPGVVLKDLGPPMVARLVREYPRADRVMSVCLRVLYGCLWRDGLLSYAPV 199  
QY 181 GOTMYLSEAVYLNDSYDGHVVGGLQGLADGVVGLDFFRKSQELRWMPGYDYVGW 240  
DB 200 GOTMYLSEAVYLNDSYDGHVVGGLQGLADGVVGLDFFRKSQELRWMPGYDYVGW 259  
QY 241 SNHSSFGYVEMEFDFRLRAFOAMVHCNNMHTLGARLPGGVECFRRGPPAMAWEGEPM 300  
DB 260 SNHSSFGYVEMEFDFRLRAFOAMVHCNNMHTLGARLPGGVECFRRGPPAMAWEGEPM 319  
QY 301 RHNLGNLGDPRARAVSVPLGGRVAREFQCRFLFAGPWLLESEIFSDVYNNSSPALGG 360  
DB 320 RHNLGNLGDPRARAVSVPLGGRVAREFQCRFLFAGPWLLESEIFSDVYNNSSPALGG 375  
QY 361 TFPAPMPPGPPPTNFSSLELEPRGQOPVAKPEGSPTA 399  
DB 376 TFPAPMPPGPPPTNFSSLELEPRGQOPVAKPEGSPTA 414

## RESULT 4

S42621  
protein-tyrosine kinase (EC 2.7.1.112) - human  
C:Species: Homo sapiens (man)  
C:Date: 07-Sep-1994 #sequence\_revision 10-Nov-1995 #text\_change 17-Mar-1999  
C:Accession: S42621  
R:Karr, T.; Holtrich, U.; Braeuninger, A.; Boehme, B.; Wolf, G.; Ruebsamen-Waigmann,  
Oncogene 8, 3433-3440, 1993  
A:Title: Structure, expression and chromosomal mapping of TKT from man and mouse: a n  
A:Reference number: S42621; MUID:94067796  
A:Accession: S42621  
A:Molecule type: mRNA  
A:Residues: 1-855 <KAR>  
A:Cross-references: EMBL:X74764; NID:g433337; PID:g433338  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; discoidin I amino-  
C:Keywords: phosphotransferase; tyrosine-specific protein kinase  
F:29-185/Domain: discoidin I amino-terminal homology <DN1>  
F:561-855/Domain: protein kinase homology <KIN>  
F:569-577/Region: protein kinase ATP-binding motif

Query Match 46.1%; Score 1002.5; DB 2; Length 855;  
Best Local Similarity 53.6%; Pred. No. 8.1e-78;  
Matches 196; Conservative 54; Mismatches 109; Indels 7; Gaps 6;

QY 5 KGFDPAPKCRYALGMDRTIPDSISASSWSSTAARHSRLSSDGDGAWCPAGSVPK 64  
DB 22 KAQVNPACRYPLGSGGQIPDEDITASSOWSESTAAYGRDSEEGDGAACPEIPVEP 81  
QY 65 E-EYLVQDLQRLHLVALVGTQGRHAGGLGKEFSRSLRYSRDRGRMMGWKDRWGQEV 123  
DB 82 DLKEFLQIDLHTLHTITLVGTQGRHAGGHGIEFAPMYKINYSRDRGTRWISWNRHKGQVL 141  
QY 124 SGNEDEGPGVVLKDLGPPMVARLVREYPRADRVMSVCLRVLYGCLWRDGLLSYAPVQ 183  
DB 142 DGSNSNPYDFLDLLEPIVARVRIPIVTDHSMVMCMRVLYGCVLMDGLSYNAPAQ 201  
QY 184 MYL--SEAVYLNDSYDGHVVGGLQGLADGVVGLDFFRKSQELRWMPGYDYVGW 241  
DB 202 FVLPGSIIYLNDSYVDG-AVGYSMTEGLGQLTDGVSGLDDFTQTQHEHYVMPGYDYV 260  
QY 242 NUSFSSGYVEMEFDFRLRAFOAMVHCNNMHTLGARLPGGVECFRRGPPAMAWEGEPM 301  
DB 242 NUSFSSGYVEMEFDFRLRAFOAMVHCNNMHTLGARLPGGVECFRRGPPAMAWEGEPM 301



Db 261 NESATNGYIEIMFEFDRINFTMTKVNMMFAKGVKIFKVCQYF-RSEASEWEPNPAIS 319  
QY 302 HNLGGNLDPPARAVSVPLGGVRAREFLOCFAGPWLFLSEIFSD-VVNNSSPALGG 360  
Db 320 FPLVLDVNPASREFTVPLHHRMASAIKQYHFDATWMMFSEITFQSDAAMYNSEAL-P 378  
QY 361 TFPAP 366  
Db 379 TSPMAP 384

RESULT 5  
I48859  
tyro 10 receptor kinase - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 19-Dec-1997  
C:Accession: I48859; 842796  
R:Lai, C.; Lemke, G.  
Oncogene 9, 877-883, 1994  
A:Title: Structure and expression of the Tyro 10 receptor tyrosine kinase.  
A:Reference number: I48859; MUID:94151011  
A:Accession: I48859  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-819 <RES>  
A:Cross-references: EMBL:X76505; NID:g435161; PID:g435162  
C:Genetics:  
A:Gene: tyro 10  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
C:Keywords: ATP  
F:526-819/Domain: protein kinase homology <KIN>  
F:534-542/Region: protein kinase ATP-binding motif

Query Match 43.8%; Score 952.5; DB 2; Length 819;  
Best Local Similarity 49.7%; Pred. No. 1.4e-73;  
Matches 189; Conservative 57; Mismatches 107; Indels 27; Gaps 7;

QY 24 IPDSISASSWSDSSTAARHSRLSSDGDGAWCPAGSVFPKE-EYQLQVDLQRLHLVALV 82  
Db 6 IPDEDITASSQWSESTAAYKGRLDSEEGDGAWCPETPQPDLDKEFLQIDLRLHFLIV 65

QY 83 GTQGRAGGLGFEFSYRLYSRGRMMGWKDRWGQVVISGNEDEPGVVLKDLGPPMV 142  
Db 66 GTQGRAGGHHGFEFAPMYKINYSRGRSWISWRNHRGKQVLDGNSNPYDFLKLLEPIV 125

QY 143 ARLVRYPRADRVMSVCLRVLYGCLWRDGLLSYAPYQOTWYL--SEAVYLNDSTYDGH 200  
Db 126 ARFVRLPVTDSHMVCMRVLYGCVLDGLVSYNAPAGQVFLPGGSIYIYLNDSVYDG- 185

QY 201 TVGGLOYGGLQADGVVGLDFFRKSQLRVWPGYDYVGSNHSFSSGVVMEFEFFDLRL 260  
Db 185 AVGYSMTEGLQITDGVSLDDFTQTHEYHVWPGYDYVGSNHSATNGFIEMFEEDRIR 244

QY 261 AFQAMQVHCNNHTLIGARLPGGVECFRGRGPAMAGEPEMRINLGNLGDPRARAVSPL 320  
Db 245 NETTMKVHCNNMFAGVKIFKVCQYF-RSEASEWEPNPAIS-SEAREVTVP 303

QY 321 GGVARFLOCFELFAGPWLFLSEIFSD--VVNNSSPALGTFPPAPWPPGPPPTNFS 378  
Db 304 HHRMASAIKQYHFDATWMMFSEITFQSDAAMYNSEAL-P 347

QY 379 SLELEPRGQOPVAKPEGSPT 398  
Db 347 ---MAPTTYDPMKLVDSNT 363

RESULT 6  
JQ0948  
A5 antigen precursor - African clawed frog  
C:Species: Xenopus laevis (African clawed frog)  
C>Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 20-Mar-1998  
C:Accession: JH0466; JQ0948

R:Takagi, S.; Hirata, T.; Agata, K.; Mochii, M.; Eguchi, G.; Fujisawa, H.  
Neuron 7, 295-307, 1991  
A:Title: The A5 antigen, a candidate for the neuronal recognition molecule, has homol  
A:Reference number: JH0466; MUID:91337458  
A:Accession: JH0466  
A:Molecule type: mRNA  
A:Residues: 1-927 <TAK>  
A:Cross-references: GB:D010467; GB:D01077; NID:g222962; PID:d1001730; PID:g222963  
A:Experimental source: tadpole, brain  
A:Note: This protein has motifs homologous to complement components C1r and C1s and t  
C:Comment: This protein is a neuronal cell surface molecule involved in the neuronal  
C:Superfamily: Xenopus A5 antigen; C1r/C1s repeat homology; discoidin I amino-termina  
C:Keywords: duplication; glycoprotein; transmembrane protein  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-927/Product: A5 antigen #status predicted <A5A>  
F:27-138/Domain: C1r/C1s repeat homology <C1R1>  
F:147-262/Domain: C1r/C1s repeat homology <C1R2>  
F:274-424/Domain: discoidin I amino-terminal homology <DN1>  
F:430-584/Domain: discoidin I amino-terminal homology <DN2>  
F:646-812/Domain: MAM homology <MAM>  
F:861-883/Domain: transmembrane #status predicted <TM>  
F:150,261,300,523,844/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.8%; Score 213; DB 2; Length 927;  
Best Local Similarity 33.3%; Pred. No. 2.9e-10;  
Matches 60; Conservative 25; Mismatches 69; Indels 26; Gaps 6;

QY 12 KCRYALGMQDRTTPDSDISASSWSDSSTAARHSRLSSDGDGAWCPAGSVFPKEEYQLQV 71  
Db 274 QCKEALGMSEGHFDQISVSSQYSMNWSAERSRLNV--ENGWTPGEDTV---KEWIOV 328

QY 72 DLQRLHLVALVGTQGRHAGGLGKE---FSRSYRLYSRDRMMGWKDRWGQVVISGN 126  
Db 329 DLENLRVSGIGIQ---GAISRETKKKFKVKSIVDSNGEDWTLKDNKHLVFTGN 384

QY 127 EDPEGVVLKDLGPPMVARLVRYPRADRVMSVCLRVLYGCLWRD-----GLIS 175  
Db 385 TDATDVVYRPFSPVITRFVRLRP-VTWENGISLRFLYGCKITDYPCSRMLGMVSGSLIS 443

RESULT 7  
JC5256  
adipocyte transcription factor, AEBP1 - Human  
C:Species: Homo sapiens (man)  
C>Date: 04-Dec-1998 #sequence\_revision 04-Dec-1998 #text\_change 04-Dec-1998  
C:Accession: JC5256  
R:Ohno, I.; Hashimoto, J.; Shimizu, K.; Takaoaka, K.; Ochi, T.; Matsubara, K.; Okubo,  
Biochem. Biophys. Res. Commun. 228, 411-414, 1996  
A:Title: A cDNA cloning of human AEBP1 from primary cultured osteoblasts and its expr  
A:Reference number: JC5256; MUID:97079196  
A:Accession: JC5256  
A:Status: preliminary  
A:Residues: 1-845 <OHN>  
A:Cross-references: DDBJ:D86479

Query Match 9.7%; Score 211.5; DB 3; Length 845;  
Best Local Similarity 24.3%; Pred. No. 3.5e-10;  
Matches 80; Conservative 54; Mismatches 144; Indels 51; Gaps 12;

QY 9 DPACRYALGMQDRTTPDSDISASSWSDSSTAARHSRLSSDGD-----DGAWCPAGSV 61  
Db 68 EKVYKCP-PIGMESHRIEDNQIRASSMLRHCLGAGQRLNMQTGATEDDYDGAWCAEDDA 126

QY 62 FPKEEYQLQVDLQRLHLVALVGTQGRHAGGLGKEFSRSYRLYSRDRMMGWKDRWGQ 121  
Db 127 ---RTQWIEVDTRTRTFTGVTQGRDS-SIHDDFTVTFVGFSGNSQTWVMYNGYEEM 182

QY 122 VISGNEDEPGVVLKDLGPPMVARLVRYPRADRVMSVCLRVLYGCLWRDGLLSYAPV 181  
Db 183 TFGHNVKDQDPVLSELPVAVFIRIYPLTWN-GSLCMRLEVIGC-----SVAPYV 233



F:1740-1741/Cleavage site: Arg-Ala (coagulation factor Xa) #status experimental  
F:2193-2345/Disulfide bonds: #status predicted

Query Match 8.8% Score 192; DB 1; Length 2351;  
Best Local Similarity 31.88; Pred. No. 5.9e-08;  
Matches 54; Conservative 32; Mismatches 68; Indels 16; Gaps 5;

Qy 9 DPAKRYALGMDRTIPDSISASS-----SWSDDTAARHSRLSSDGDGAWCPAGSV 61  
Db 2189 DLNCSMPLGMSKAISDAITASSYFTNMFATWSPSKARHLQGRSN-----AWRFQVN- 2244  
Qy 62 FPKKEEYLVLDLRLHLVALVGTQGRHAGGLGKEFSRYRLYSRDGRRWMGKDRWQGE 121  
Db 2244 --NPKEWLVDFQTKMKTGVTGQVKS--LLTSWYVKEFLISSQDGHQWTLFFQNGKVK 2300  
Qy 122 VISGNEDEPGVWLKDLGPPVARLVREYPRADRVMSVCLRVLYGCLWRD 171  
Db 2301 VFQGNQDSFTPVNSLDPLTLRYLRHPOS--WVHQIALRMEVLGCEAQD 2349

RESULT 9  
A44258  
factor VIII-associated gene B hypothetical protein - human  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 31-Oct-1997  
C:Accession: A44258  
R:Levinson, B.; Kenwright, S.; Gamel, P.; Fisher, K.; Gitschler, J.  
Genomics 14, 585-589, 1992  
A:Title: Evidence for a third transcript from the human factor VIII gene.  
A:Reference number: A44258; MUID:93052386  
A:Accession: A44258  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-216 <LEV>  
A:Cross-references: GB:M90707; NID:g182316; PID:g182317  
A:Superfamily: coagulation factor VIII; discoidin I amino-terminal homology; ferroxidase  
F:1-53/Domain: discoidin I amino-terminal homology #status atypical <DN1>  
F:57-210/Domain: discoidin I amino-terminal homology <DN2>

Query Match 8.8% Score 192; DB 2; Length 216;  
Best Local Similarity 31.88; Pred. No. 2.8e-09;  
Matches 54; Conservative 32; Mismatches 68; Indels 16; Gaps 5;

Qy 9 DPAKRYALGMDRTIPDSISASS-----SWSDDTAARHSRLSSDGDGAWCPAGSV 61  
Db 54 DLNCSMPLGMSKAISDAITASSYFTNMFATWSPSKARHLQGRSN-----AWRFQVN- 109  
Qy 62 FPKKEEYLVLDLRLHLVALVGTQGRHAGGLGKEFSRYRLYSRDGRRWMGKDRWQGE 121  
Db 109 --NPKEWLVDFQTKMKTGVTGQVKS--LLTSWYVKEFLISSQDGHQWTLFFQNGKVK 165  
Qy 122 VISGNEDEPGVWLKDLGPPVARLVREYPRADRVMSVCLRVLYGCLWRD 171  
Db 166 VFQGNQDSFTPVNSLDPLTLRYLRHPOS--WVHQIALRMEVLGCEAQD 214

RESULT 10  
KFHUS  
coagulation factor V precursor - human  
N:Alternate names: coagulation labile factor; proaccelerin  
C:Species: Homo sapiens (man)  
C:Date: 19-May-1989 #sequence\_revision 02-Jun-1995 #text\_change 31-Oct-1997  
C:Accession: A56172; A42344; A28028; A27498; A25897  
R:Cripe, L.D.; Moore, K.D.; Kane, W.H.  
Biochemistry 31, 3777-3785, 1992  
A:Title: Structure of the gene for human coagulation factor V.  
A:Reference number: A42344; MUID:92232668  
A:Accession: A56172  
A:Molecule type: DNA  
A:Residues: 1-2224 <CR>  
A:Cross-references: GB:J05368

A:Accession: A42344  
A:Molecule type: DNA  
A:Residues: 48-58-79-89;120-130;191-201;239-249;313-323;368-378;428-437;461-471;533-5  
2070;2111-2120;2172-2181 <CR2>  
R:Jenny, R.J.; Pittman, D.D.; Toole, J.J.; Kriz, R.W.; Aldape, R.A.; Hewick, R.M.; Ka  
Proc. Natl. Acad. Sci. U.S.A. 84, 4846-4850, 1987  
A:Title: Complete cDNA and derived amino acid sequence of human factor V.  
A:Reference number: A28028; MUID:87260886  
A:Accession: A28028  
A:Molecule type: mRNA  
A:Residues: 1-857, 'R', 859-864, 'R', 866-924, 'E', 926-1763, 'I', 1765-2212, 'T', 2214-2224 <J  
A:Cross-references: GB:M16967  
A:Note: parts of this sequence, including the amino end of the mature protein, were d  
R:Kane, W.H.; Ichinose, A.; Hagen, F.S.; Davie, E.W.  
Biochemistry 26, 6508-6514, 1987  
A:Title: Cloning of cDNAs coding for the heavy chain region and connecting region of  
A:Reference number: A27498; MUID:88107560  
A:Accession: A27498  
A:Molecule type: mRNA  
A:Residues: 1-1284, 'I', 1286-1600 <KAN>  
A:Cross-references: GB:M17785  
A:Note: parts of this sequence were determined by protein sequencing  
R:Kane, W.H.; Davie, E.W.  
Proc. Natl. Acad. Sci. U.S.A. 83, 6800-6804, 1986  
A:Title: Cloning of a cDNA coding for human factor V, a blood coagulation factor homo  
A:Reference number: A25897; MUID:86313665  
A:Accession: A25897  
A:Molecule type: mRNA  
A:Residues: 1188-1215, 1315-2224 <KA2>  
A:Cross-references: GB:M14335  
A:Note: parts of this sequence were determined by protein sequencing  
R:Keller, F.G.; Ortel, T.L.; Quinn-Alten, M.A.; Kane, W.H.  
Biochemistry 34, 4118-4124, 1995  
A:Title: Thrombin-catalyzed activation of recombinant human factor V.  
A:Reference number: A56139; MUID:95210278  
A:Contents: annotation; thrombin cleavage sites  
C:Comment: Factor V is activated by thrombin and partially by coagulation factor Xa.  
C:Genetics:  
A:Gene: GDB:F5  
A:Cross-references: GDB:119896; OMIM:227400  
A:Map position: 1q23-1q23  
A:Introns: 53/2; 84/1; 125/1; 196/1; 244/1; 318/1; 373/2; 432/3; 466/1; 537/3; 588/1;  
C:Function:  
A:Description:  
A:Pathway: blood coagulation  
C:Superfamily: coagulation factor V; discoidin I amino-terminal homology; ferroxidase  
C:Keywords: blood coagulation; duplication; glycoprotein; phospholipid binding; plasm  
F:1-28/Domain: signal sequence #status predicted <SIG>  
F:29-2224/Product: coagulation factor V #status predicted <MAT>  
F:29-737/Product: coagulation factor Va heavy chain #status experimental <VAH>  
F:29-345/Domain: A1 <DA1>  
F:33-329/Domain: ferroxidase repeat homology <FO1>  
F:346-691/Domain: A2 <DA2>  
F:351-684/Domain: ferroxidase repeat homology <FO2>  
F:692-1573/Domain: B <DOB>  
F:1183-1461/Region: 9-residue repeats (O-X-T/N-L-S-P-D-L-S)  
F:1574-2224/Product: coagulation factor Va light chain #status experimental <VAL>  
F:1574-1905/Domain: A3 <DA3>  
F:1581-1905/Domain: ferroxidase repeat homology <FO3>  
F:1667-1765/Region: phospholipid binding #status predicted  
F:1906-2064/Domain: C1 <DC1>  
F:1906-2061/Domain: discoidin I amino-terminal homology <DN1>  
F:2065-2224/Domain: C2 <DC2>  
F:2065-2221/Domain: discoidin I amino-terminal homology <DN2>  
F:51-55,239,297,460,468,554,741,752,760,776,782,821,938,977,1074,1083,1103,1106,1479,  
F:167-193,248-329,500-526,603-684,1725-1751,1907-2061,2066-2221/Disulfide bonds: #sta  
F:334-335/Cleavage site: Arg-Asn (protein C) #status predicted  
F:363,693,1546/Binding site: sulfate (Tyr) (covalent) #status predicted  
F:376-377/Cleavage site: Arg-Ser (coagulation factor Xa) #status predicted  
F:382,1338/Binding site: carboxylate (Asn) (covalent) #status absent  
F:534-535/Cleavage site: Arg-Gly (protein C) #status predicted  
F:737-738/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status experiment  
F:1046-1047/Cleavage site: Arg-Thr (coagulation factor Xa, thrombin) #status experime

A:Cross-references: GB:L05573; NID:gl92456; PID:gl92457  
C:Superfamily: coagulation factor VII; discoidin I amino-terminal homology  
F;1-19/Domain: signal sequence #status predicted <Sig>

F;1893-2051/Domain: C1 <DC1>

1-19/domain: signal sequence #status predicted <SIG>

F:2052-2211/Domain: C2 <DC2>  
F:2052-2208/Domain: discoidin I amino-terminal homology <DN2>  
F:167-193\_248-329,499-525/disulfide bonds: #status experimental  
F:225-239,297\_382,460,553,587,745,756,774,780,902,952,964,1044,1053,1062,1071,1078,1094,  
F:324-335/Cleavage site: Arg-Asn (protein C) #status predicted  
F:363\_697\_1537/Binding site: sulfate (Tyr) (covalent) #status predicted  
F:376-377/Cleavage site: Arg-Ser (coagulation factor Xa) #status predicted  
F:533-534/Cleavage site: Arg-Gly (protein C) #status predicted  
F:607-688,1712-1738,1894-2048,2053-2208/disulfide bonds: #status predicted  
F:741-742/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status predicted  
F:1034-1035/Cleavage site: Arg-Ser (coagulation factor Xa, thrombin) #status predicted  
F:1564-1565/Cleavage site: Arg-Ser (thrombin) #status experimental

Query Match 8.0%; Score 174.5; DB 1; Length 2211;  
Best Local Similarity 27.8%; Pred. No. 1.7e-06;  
Matches 45; Conservative 32; Mismatches 72; Indels 13; Gaps 6;

Qy	13 CRYALGMDRTIPDSDISASS---WSDSTAARHSRLSSDDGDGAWCAGSVFPKEEY 68       : :   :   :   :   :   :   :   :
Db	2053 CSTPLGMESGKTENKQTASSFKKSWGNYWEPIARLNAGQRVNAQAKAN---NNNOW 2109

```

Qy 69 LQVDLQRLHLVALVGTQGRHAGGKKE-FRSYRLRYSDGRRRWGWKDRWG--QEVISS 125
    ||:||::: ||| | | :||: || | | :||: |||
Db 2110 LQIDLKIKKIPAVTQG--CKSLSEMVYKSYTHYSDGTDWPKYREKSSWDRKIFEG 2167

```

**Qy**      126 NEDPEGVVLKDLGPPMVARLVRFPRADRVMSVCURVELYGC 167  
         | : | | ||::| : | : | : ||::||  
**Dd**      2168 NNNVRGHVKFNFPPIIGRIFRIIPKTWN-QSIALRLLEFGC 2208

RESULT 14  
A36479

milk rat globule membrane protein - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 28-Mar-1991 #sequence\_revision 28-Mar-1991 #text\_change 07-Aug-1998  
C:Accession: A36479

A: Title: cDNA cloning of a mouse mammary epithelial cell surface protein receptor.  
A: Reference number: A36479. MUID: 91046008

A;Accession: A36479  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-463 /5min/

A; Cross-references: GB:M38337; NID:g199142; PID:g199143  
C; Superfamily: milk fat globule protein; discoidin I amino-terminal homology  
C; Keywords: membrane protein

F;68-107/Domain: EGF homology <EGF>  
F;147-303/Domain: discoidin I amino-terminal homology <DN1>  
F;307-463/Domain: discoidin I amino-terminal homology <DN2>

Query Match 8.0%; Score 174.5; DB 2; Length 463;  
Best Local Similarity 32.9%; Pred. No. 2.3e-07;

Q7 13 CRYALGMDRTIPDSDISASSS---WSDSTAARHSRLESDDGD---AWCPAGSVTPKEE 66

QY 67 EYLOVDLQRLHLVALVGTGQ - RHAGGLGKFEFSRYRLRYSRDGRRRWGKDWGQEVISG 125

Qy 126 NEDPEGVVLKDLGPPMVARLVRFYPR--DRVMSVCLRLVELYGC 167

DD 423 NLDNNSHANN1FEAFFMARIKRVJFVSWHNR---ITLRLELLGC 463

RESULT 15

ags protein precursor - rat  
N;Alternate names: O-acetyl-Gd3 ganglioside  
C;Species: Rattus norvegicus (Norway rat)  
C;date: 26-Sep-1996 #sequence\_revision 01-Nov-1996 #text\_change 13-Nov-1998  
C;Accession: JCN4915  
R;Ogura, K.; Nara, K.; Watanabe, Y.; Kohno, K.; Tai, T.; Sanai, Y.  
Biochem. Biophys. Res. Commun. 225, 932-938, 1996  
A;Title: Cloning and expression of cDNA for O-acetylation of GD3 ganglioside.  
A;Reference number: JCN4915; MUID:96374422

A;Accession: J04915  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-427 <OGU>

A; Cross-references: DBJQ:084088; NID:glb20006; PID:glb20007  
 A; Experimental source: CST cell  
 C; Comment: This protein is required for the O-acetylation of disialoganglioside sialic acid  
 C; Genetics:

A:Gene: ags  
C:Superfamily: milk fat globule protein; discoidin I amino-terminal homology; EGF hom  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:28-60/Domain: EGF homology <EG1>

F:68-10//Domain: EGF homology <EG2>  
F:110-267//Domain: discoidin I amino-terminal homology <DN1>  
F:271-427//Domain: discoidin I amino-terminal homology <DN2>

Query Match	7.9%	Score 172.5;	DB 2;	Length 427;
Best Local Similarity	32.9%;	pred. No. 3.1e-07;		
Matches	54:	Conservative	27:	Mismatches 66;
				Indels 17;
				Gaps 7;

QY 13 CRYALGMQDRITIPDSIDASSS---WSDSTAARHSRLSSDGDG---AWCPAGSVFPKEE 66  
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
Db 272 CSBPGLGNKNTTIPDSOTIATSSVKTKNTPAFGVDPHLCPI DNOCKTNATPACSN---SAF 328

**OY**   67 EYLQVDLQRLLHVALVYGTOG-RHAGGLGEFSRYRLYSRDGRWGWKDRWGDEIVSG 125  
| : | | | |  
**DH**   328 EWLVDTGTOTAYCTTTCATAPDSCHI--GVVASVKVAHSDDCVMTVSFSCSTCKVFQC 296  
| : | | | |

QY 126 NEDPEGVVLKDLGPPMVARLVREVPRA--DRVMSVCLRVELYGC 167

397 NTDNNGCHVETETETEDVAPVWVMDYGVND--TMDVETVCC 427

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

Job time: 7588 sec



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 4, 1999, 01:52:00 ; Search time 18.36 Seconds  
(without alignments)  
248.367 Million cell updates/sec

Title: US-08-170-558-8  
Perfect score: 2175  
Sequence: 1 DADMKGHDPKCRYALGMQ.....LELEPRGQPPVAKPEGSPTA 399

Scoring table: BLOSUM62

Searched: 119832 seqs, 11428610 residues

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/PCTUS9\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2175	100.0	913	1	US-08-445-640-4
2	2175	100.0	399	1	US-08-445-640-8
3	2167	99.6	919	1	US-08-336-343A-2
4	1002.5	46.1	855	1	US-08-336-343A-4
5	995.5	45.8	854	2	US-08-456-647B-20
6	995.5	45.8	854	2	US-08-237-401A-20
7	210	9.7	2133	2	US-08-670-707A-37
8	210	9.7	1443	2	US-08-670-707A-39
9	201.5	9.3	1128	1	US-08-111-939-2
10	193	8.9	218	1	US-07-607-538C-5
11	192	8.8	2332	1	US-07-864-004B-4
12	192	8.8	2351	1	US-08-121-202-2
13	192	8.8	2332	1	US-08-251-937A-4
14	192	8.8	2332	1	US-08-212-133A-2
15	192	8.8	2351	1	US-08-366-831A-2
16	192	8.8	2332	1	US-08-276-594A-2
17	192	8.8	1471	1	US-08-683-839B-3
18	192	8.8	2332	1	US-08-474-503-2
19	192	8.8	2332	2	US-08-670-707A-2
20	192	8.8	1661	2	US-08-882-083-2
21	192	8.8	1661	2	US-08-558-107-2
22	192	8.8	2332	3	PCT-US93-03275-4
23	192	8.8	2332	3	PCT-US94-13200-2
24	179.5	8.3	218	1	US-07-607-538C-4
25	177.5	8.2	218	1	US-07-607-538C-2
26	177.5	8.2	217	1	US-07-607-538C-3
27	175	8.0	2319	1	US-08-212-133A-8
28	175	8.0	2319	1	US-08-474-503-6
29	175	8.0	2319	2	US-08-670-707A-6
30	175	8.0	2319	3	PCT-US94-13200-6
31	171	7.9	480	2	US-08-480-229C-10
32	171	7.9	480	2	US-08-659-235C-10
33	166.5	7.7	513	2	US-08-480-229C-14
34	166.5	7.7	513	2	US-08-659-235C-14
35	163.5	7.5	320	2	US-08-480-229C-20
36	163.5	7.5	320	2	US-08-659-235C-20
37	158.5	7.3	321	2	US-08-480-229C-21
38	158.5	7.3	321	2	US-08-659-235C-21
39	147.5	6.8	109	1	US-08-111-939-19

Sequence 23, Appl  
Sequence 21, Appl  
Sequence 25, Appl  
Sequence 20, Appl  
Sequence 24, Appl  
Sequence 3, Appl

## ALIGNMENTS

RESULT 1  
US-08-445-640-4  
; Sequence 4, Application US/08445640  
; Patent No. 5709858  
; GENERAL INFORMATION:  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Mark, Melanie R.  
; APPLICANT: Scadden, David T.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Baron, Will F.  
; TITLE OF INVENTION: Protein Tyrosine Kinases  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/445,640  
; FILING DATE: 22-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/170558  
; FILING DATE: 20-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/157563  
; FILING DATE: 23-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hasak, Janet E.  
; REGISTRATION NUMBER: 28,616  
; REFERENCE/DOCKET NUMBER: 854C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-1896  
; TELEFAX: 415/942-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 913 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; US-08-445-640-4

Query Match 100.0%; Score 2175; DB 1; Length 913;  
Best Local Similarity 100.0%; Pred No. 3.3e-216;  
Matches 399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DADMKGHDPKCRYALGMQDRTIPDSISASSSWSDSTAARHSRLSSDGDGAWCPAGS 60  
Db 19 DADMKGHDPKCRYALGMQDRTIPDSISASSSWSDSTAARHSRLSSDGDGAWCPAGS 78  
QY 61 VFPEEYLOVDLQRLHLVALVGTQGRHAGGLCKEFSRSLRYLRSDGRRWGKDRWQ 120  
Db 79 VFPEEYLOVDLQRLHLVALVGTQGRHAGGLCKEFSRSLRYLRSDGRRWGKDRWQ 138

QY 121 EVISGNEDEPGVVLKDLGPPMVARLVRFYPRADRVMSVCLRVLYGCLWRDGLLSYAPV 180  
DB 139 EVISGNEDEPGVVLKDLGPPMVARLVRFYPRADRVMSVCLRVLYGCLWRDGLLSYAPV 198  
QY 181 GQTYLSEAVYLNDSYDGHVTGGVGLQGLGQLADGVVGLDDFRKSOELRVWPGYDYVGW 240  
DB 199 GQTYLSEAVYLNDSYDGHVTGGVGLQGLGQLADGVVGLDDFRKSOELRVWPGYDYVGW 258  
QY 241 SNHFSFGVYVMEFEFDRLRAFAQMVHCNNMHTLGARLPGGVCECRPRRGPAMAWEGEPM 300  
DB 259 SNHFSFGVYVMEFEFDRLRAFAQMVHCNNMHTLGARLPGGVCECRPRRGPAMAWEGEPM 318  
QY 301 RHNLGNLGDPRARAVSVPLGGRVARELQCRFLFAGPWLLEFSEISFISDVVNNSSPALGG 360  
DB 319 RHNLGNLGDPRARAVSVPLGGRVARELQCRFLFAGPWLLEFSEISFISDVVNNSSPALGG 378  
QY 361 TFPAPMPPGPPPTNFSSLELEPRGQOPVAKPEGSPTA 399  
DB 379 TFPAPMPPGPPPTNFSSLELEPRGQOPVAKPEGSPTA 417

## RESULT 2

US-08-445-640-8  
; Sequence 8, Application US/08445640  
; Patent No. 5709858

## GENERAL INFORMATION:

APPLICANT: Godowski, Paul J.  
APPLICANT: Mark, Melanie R.  
APPLICANT: Scadden, David T.  
APPLICANT: Baker, Kevin P.  
APPLICANT: Baron, Will F.  
TITLE OF INVENTION: Protein Tyrosine Kinases  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESS: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080

## COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25 inch, 360 kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/445,640  
FILING DATE: 22-MAY-1995

## CLASSIFICATION: 435

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/170558

## FILING DATE: 20-DEC-1993

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/157563

## FILING DATE: 23-NOV-1993

## ATTORNEY/AGENT INFORMATION:

NAME: Hasak, Janet E.

REGISTRATION NUMBER: 28,616

REFERENCE/DOCKET NUMBER: 854C2

## TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-1896

TELEFAX: 415/952-9881

TELEX: 910/371-7168

## INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 399 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-445-640-8

Query Match 100.0%; Score 2175; DB 1; Length 399;  
Best Local Similarity 100.0%; Pred. No. 1e-216;

Matches 399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DADMKGHFDPAKCRVALGMDQRTIPDSISASSSSWSDSTAARHSRLSSDGDGAWCPAGS 60  
DB 1 DADMKGHFDPAKCRVALGMDQRTIPDSISASSSSWSDSTAARHSRLSSDGDGAWCPAGS 60  
QY 61 VFPKKEEYLQVDLQRLHLVALVGTGGRHAGGLGKEFSRSYRLRYSRDGRRWGKDRWGQ 120  
DB 61 VFPKKEEYLQVDLQRLHLVALVGTGGRHAGGLGKEFSRSYRLRYSRDGRRWGKDRWGQ 120  
QY 121 EVISGNEDEPGVVLKDLGPPMVARLVRFYPRADRVMSVCLRVLYGCLWRDGLLSYAPV 180  
DB 121 EVISGNEDEPGVVLKDLGPPMVARLVRFYPRADRVMSVCLRVLYGCLWRDGLLSYAPV 180  
QY 181 GQTYLSEAVYLNDSYDGHVTGGVGLQGLGQLADGVVGLDDFRKSOELRVWPGYDYVGW 240  
DB 181 GQTYLSEAVYLNDSYDGHVTGGVGLQGLGQLADGVVGLDDFRKSOELRVWPGYDYVGW 240  
QY 241 SNHFSFGVYVMEFEFDRLRAFAQMVHCNNMHTLGARLPGGVCECRPRRGPAMAWEGEPM 300  
DB 241 SNHFSFGVYVMEFEFDRLRAFAQMVHCNNMHTLGARLPGGVCECRPRRGPAMAWEGEPM 300  
QY 301 RHNLGNLGDPRARAVSVPLGGRVARELQCRFLFAGPWLLEFSEISFISDVVNNSSPALGG 360  
DB 301 RHNLGNLGDPRARAVSVPLGGRVARELQCRFLFAGPWLLEFSEISFISDVVNNSSPALGG 360  
QY 361 TFPAPMPPGPPPTNFSSLELEPRGQOPVAKPEGSPTA 399  
DB 361 TFPAPMPPGPPPTNFSSLELEPRGQOPVAKPEGSPTA 399

RESULT 3  
US-08-336-343A-2  
; Sequence 2, Application US/08336343A  
; Patent No. 5677144  
GENERAL INFORMATION:  
APPLICANT: Ullrich, Axel  
APPLICANT: Alves, Frauke  
TITLE OF INVENTION: CCK-2, A No. 5677144el Receptor Tyrosine Kinase  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/336,343A  
FILING DATE: 08-NOV-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7683-065  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 790-9090  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 919 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-336-343A-2



Query Match 99.6%; Score 2167; DB 1; Length 919;

Best Local Similarity 99.7%; Pred. No. 2.2e-215;

Matches 398; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DADMKGHFPAKCRYALGMQDRTPDSDISASSSSWSDSTAARHSRLSSDGDGAWCPAGS 60  
DB 19 DADMKGHFPAKCRYALGMQDRTPDSDISASSSSWSDSTAARHSRLSSDGDGAWCPAGS 78  
QY 61 VPKBEYEQVDLQRLHVALVGTQGRHAGGLGKEFSRSYRLRYSRDGRWGWKDRWGQ 120  
DB 79 VPKBEYEQVDLQRLHVALVGTQGRHAGGLGKEFSRSYRLRYSRDGRWGWKDRWGQ 138  
QY 121 EVLSGNEDEGGVVLKDLGPPMVARLRFYPRADRVMSVCLRVLYGCLWRDGLLSYTA 180  
DB 139 EVLSGNEDEGGVVLKDLGPPMVARLRFYPRADRVMSVCLRVLYGCLWRDGLLSYTA 198  
QY 181 GQPMYLSAAYLNDSTYDGTGGVGLQGLADGVGLDFFRKSQELRVWPGDYDYGW 240  
DB 199 GQPMYLSAAYLNDSTYDGTGGVGLQGLADGVGLDFFRKSQELRVWPGDYDYGW 258  
QY 241 SNHSFSGVYEMEFEDRLRAFOAMQVHCNNMHTLGARLPGGVCECRFRGPPAMAWEGEPM 300  
DB 259 SNHSFSGVYEMEFEDRLRAFOAMQVHCNNMHTLGARLPGGVCECRFRGPPAMAWEGEPM 318  
QY 301 RNLGNLGDPRARAVSVPLGGVAREFLQCRFLFAGPWLLEISFISDVVNSSPALGG 360  
DB 319 RNLGNLGDPRARAVSVPLGGVAREFLQCRFLFAGPWLLEISFISDVVNSSPALGG 378  
QY 361 TFPAPWPPGPPPTNFSSLEPRGQQPVAKPEGSPTA 399  
DB 379 TFPAPWPPGPPPTNFSSLEPRGQQPVAKPEGSPTA 417

## RESULT 4

US-08-336-343A-4

; Sequence 4, Application US/08336343A

; Patent No. 5677144

; GENERAL INFORMATION:

; APPLICANT: Ullrich, Axel

; APPLICANT: Alves, Frauke

; TITLE OF INVENTION: CCK-2, A No. 5677144el Receptor Tyrosine Kinase

; NUMBER OF SEQUENCES: 43

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10036-2711

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/336,343A

; FILING DATE: 08-NOV-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Coruzzi, Laura A.

; REGISTRATION NUMBER: 30,742

; REFERENCE/DOCKET NUMBER: 7683-065

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 790-9090

; TELEFAX: (212) 869-9741/8864

; TELEX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 855 amino acids

; TYPE: amino acid

; TOPOLOGY: unknown

; MOLECULE TYPE: protein

; US-08-336-343A-4

Query Match 46.1%; Score 1002.5; DB 1; Length 855;

Best Local Similarity 53.6%; Pred. No. 4.2e-95;

Matches 196; Conservative 54; Mismatches 109; Indels 7; Gaps 6;

QY 5 KGHDPKAKRYALGMQDRTPDSDISASSSSWSDSTAARHSRLSSDGDGAWCPAGSVFPK 64  
DB 22 KAQYNPAICRYPLGMSGGQIPDEDITASSQWSESTAARKYGLRDLSEEDGAWCPEIPVED 81  
QY 65 E-EYLOVDLQRLHVALVGTQGRHAGGLGKEFSRSYRLRYSRDGRWGWKDRWGQEV 123  
DB 82 DLKFLQIDLHVALVGTQGRHAGGLGKEFSRSYRLRYSRDGRWGWKDRWGQEV 141  
QY 124 SGNEDPEGVVLKDLGPPMVARLRFYPRADRVMSVCLRVLYGCLWRDGLLSYTA 183  
DB 142 DGSNPYDIFLKDLEPPIVARFVRFIPVTDHSMNVMCRVELYGCWLDGLVSNAPAGQ 201  
QY 184 MYL--SEAVYLNSTYDGTGGVGLQGLADGVGLDFFRKSQELRVWPGDYDYGW 241  
DB 202 FVLPGGSIILYLNDSYDVG-AVGYSMTGLGQLTDGVSGLDDFTQTHEYHWPGDYDYGW 260  
QY 242 NHFSFGVYEMEFEDRLRAFOAMQVHCNNMHTLGARLPGGVCECRFRGPPAMAWEGEPM 301  
DB 261 NESATNGIIEIMFEDRIRNFTTMKVHCNNMFAKVIFKEVOCYF-RSEASEWEPNAIS 319  
QY 302 HNLGNLGDPRARAVSVPLGGVAREFLQCRFLFAGPWLLEISFISDVVNSSPALGG 360  
DB 320 FPLVLDVNPASARVTVPLHHRMASAIKCOYHEADTWMESEITFQSDAAMYNSEAL-P 378  
QY 361 TFPAP 366  
DB 379 TSPMAP 384

## RESULT 5

US-08-456-647B-20

; Sequence 20, Application US/08456647B

; Patent No. 5811516

; GENERAL INFORMATION:

; APPLICANT: Lenke Ph.D. et al., Greg E.

; TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES

; NUMBER OF SEQUENCES: 54

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 4225 Executive Square, Suite 1400

; CITY: La Jolla

; STATE: CA

; COUNTRY: US

; ZIP: 92037

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/456,647B

; FILING DATE: 02-JUN-1995

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/237,401

; FILING DATE: 02-MAY-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/884,486

; FILING DATE: 15-MAY-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Wetherell Ph.D., John R.

; REGISTRATION NUMBER: 31,678

; REFERENCE/DOCKET NUMBER: 07251/007002

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619) 678-5070

; TELEFAX: (619) 678-5099

; INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:  
LENGTH: 854 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-456-647B-20

Query Match 45.8%; Score 995.5; DB 2; Length 854;  
Best Local Similarity 49.6%; Pred. No. 2.2e-94;  
Matches 198; Conservative 58; Mismatches 116; Indels 27; Gaps 7;

QY 5 KGHEDPAKRYALGMDRTIPDSISASSWSDSTAARHSRLSSDGDGAWCPAGSVFPK 64  
DB 22 KQVNPALCRYPGLMSGGHIPDEDITASSQWSESTAARYGLDSEEGDGCWCPPIPVQPD 81  
QY 65 E-EYQLVDLQRLHLVALVGTQGRHAGGLGKEFSRSYRLRYSRDGRWGMKDRWGQEV 123  
DB 82 DLKEFLQIDLRLTHFTILVGTQGRHAGGHIETAFAPMYKINYSRDGSRWISWRNRHGQVL 141  
QY 124 SGNEDPEGVLKDLGPPMARVRLRYPRADRVMSVCLRVLYGCLWRDGLLSYAPVQGT 183  
DB 142 DGSNPYDVFLKLEPPVAVRFVRLIPVTDHSMNVMCMRVLYGCVMLDGLVSYNAPAGQ 201  
QY 184 MYL--SEAVYLDSTYDGHVTGGLOYGLGLADGVVGLDDFRKSQLRWPGDYVWGS 241  
DB 202 FVLPGGSIYLDNSYDYG-AVGYSMTEGLGLDGTGVSLDDFTQTHEYHVMWPGDYVWGR 260  
QY 242 NHSFSSGVMEFEFDRLARFAQMVHNCNNHMTLGARLPGGVCEFRFRGPAMAWEGEPMR 301  
DB 261 NESATNGFIEMFEFDRIRNFTTKVHCNMFAGVKIFKEVQCYF-RSEASEWEPTAVY 319  
QY 302 HNLGNLGDPRARAVSVPLGGRVARFLQCRFLFAGPWLLEFSEISFISD--VYNNSSPALG 359  
DB 320 PFLVLDVNPASREVTPLHHRMASAIKQYHFADTWMESEITFQSDAAMYNNS----- 375  
QY 360 GTFFPAPWPPGPPPTNFSSLELEPRGQOPVAKPEGSP 398  
DB 375 GALPTSP-----MAPTTYDPMKVDSDNT 398

## RESULT 6

US-08-237-401A-20  
Sequence 20, Application US/08237401A  
Patent No. 5837448  
GENERAL INFORMATION:  
APPLICANT: Lemke Ph.D. et al., Greg E.  
TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES  
NUMBER OF SEQUENCES: 54  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CA  
COUNTRY: US  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/237,401A  
FILING DATE: 02-MAY-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/884,486  
FILING DATE: 15-MAY-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Haile Ph.D., Lisa A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07251/007001  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 678-5070  
TELEFAX: (619) 678-5099  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 854 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-237-401A-20

Query Match 45.8%; Score 995.5; DB 2; Length 854;  
Best Local Similarity 49.6%; Pred. No. 2.2e-94;  
Matches 198; Conservative 58; Mismatches 116; Indels 27; Gaps 7;

QY 5 KGHEDPAKRYALGMDRTIPDSISASSWSDSTAARHSRLSSDGDGAWCPAGSVFPK 64  
DB 22 KQVNPALCRYPGLMSGGHIPDEDITASSQWSESTAARYGLDSEEGDGCWCPPIPVQPD 81  
QY 65 E-EYQLVDLQRLHLVALVGTQGRHAGGLGKEFSRSYRLRYSRDGRWGMKDRWGQEV 123  
DB 82 DLKEFLQIDLRLTHFTILVGTQGRHAGGHIETAFAPMYKINYSRDGSRWISWRNRHGQVL 141  
QY 124 SGNEDPEGVLKDLGPPMARVRLRYPRADRVMSVCLRVLYGCLWRDGLLSYAPVQGT 183  
DB 142 DGSNPYDVFLKLEPPVAVRFVRLIPVTDHSMNVMCMRVLYGCVMLDGLVSYNAPAGQ 201  
QY 184 MYL--SEAVYLDSTYDGHVTGGLOYGLGLADGVVGLDDFRKSQLRWPGDYVWGS 241  
DB 202 FVLPGGSIYLDNSYDYG-AVGYSMTEGLGLDGTGVSLDDFTQTHEYHVMWPGDYVWGR 260  
QY 242 NHSFSSGVMEFEFDRLARFAQMVHNCNNHMTLGARLPGGVCEFRFRGPAMAWEGEPMR 301  
DB 261 NESATNGFIEMFEFDRIRNFTTKVHCNMFAGVKIFKEVQCYF-RSEASEWEPTAVY 319  
QY 302 HNLGNLGDPRARAVSVPLGGRVARFLQCRFLFAGPWLLEFSEISFISD--VYNNSSPALG 359  
DB 320 PFLVLDVNPASREVTPLHHRMASAIKQYHFADTWMESEITFQSDAAMYNNS----- 375  
QY 360 GTFFPAPWPPGPPPTNFSSLELEPRGQOPVAKPEGSP 398  
DB 375 GALPTSP-----MAPTTYDPMKVDSDNT 398

## RESULT 7

US-08-670-707A-37  
Sequence 37, Application US/08670707A  
Patent No. 5859204  
GENERAL INFORMATION:  
APPLICANT: Lollar, John S.  
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
STREET: 5370 Manhattan Circle Suite 201  
CITY: Boulder  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/670,707A  
FILING DATE: 26-JUN-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US94/13200  
FILING DATE: 15-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/212,133

;; FILING DATE: 11-MAR-1994  
;; PRIOR APPLICATION DATA: US 07/864,004  
;; APPLICATION NUMBER: US 07/864,004  
;; FILING DATE: 07-APR-1992  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Greenlee, Lorance L.  
;; REGISTRATION NUMBER: 27,894  
;; REFERENCE/DOCKET NUMBER: 75-95F  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 303/499-8080  
;; TELEFAX: 303/499-8089  
;; INFORMATION FOR SEQ ID NO: 37:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 2133 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-670-707A-37

Query Match 9.7%; Score 210; DB 2; Length 2133;  
Best Local Similarity 34.7%; Pred. No. 1.2e-12;  
Matches 59; Conservative 24; Mismatches 71; Indels 16; Gaps 5;  
QY 9 DPAKCRYALGMQDRTIPDSISASS-----SWSDDTAARHSRLSSDGDGAWCPAGSV 61  
Db 1971 DLNCSMPLGMQNAISDSQITASSLSNIFATWSPSQ-----RLHLOGRTNNAWRPRVS- 2026  
QY 62 FPKEEYLQVLDLQRLHLVALVGTQGRHAGGLGKEFSRSLRYLRYSRDRGRRWGWKDRWGOE 121  
Db 2026 --SAEWMQLVDLQKTVKVTGTTGVKS--LLSSMYVKEFLVSSQDGRRTWTLFLQDGHK 2082  
QY 122 VISGNEDEPGVVLKDLGPPVAVLRVRYPRADRVMSVCLRVLYGCLWRD 171  
Db 2083 VFOGNDSTPTVNVNADPLPFLTRYLRHP-TSWAQHIALRLVLCGEAOD 2131

RESULT 8  
US-08-670-707A-39  
; Sequence 39, Application US/08670707A  
; Patent No. 5859204  
; GENERAL INFORMATION:  
; APPLICANT: Lollar, John S.  
; TITLE OF INVENTION: Hybrid Human/Animal Factor VIII  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
; STREET: 5370 Manhattan Circle Suite 201  
; CITY: Boulder  
; STATE: Colorado  
; COUNTRY: USA  
; ZIP: 80303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/670,707A  
; FILING DATE: 26-JUN-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/US94/13200  
; FILING DATE: 15-NOV-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/212,133  
; FILING DATE: 11-MAR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/864,004  
; FILING DATE: 07-APR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Greenlee, Lorance L.  
; REGISTRATION NUMBER: 27,894

;; REFERENCE/DOCKET NUMBER: 75-95F  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 303/499-8080  
;; TELEFAX: 303/499-8089  
;; INFORMATION FOR SEQ ID NO: 39:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 1443 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-670-707A-39

Query Match 9.7%; Score 210; DB 2; Length 1443;  
Best Local Similarity 34.7%; Pred. No. 6.8e-13;  
Matches 59; Conservative 24; Mismatches 71; Indels 16; Gaps 5;  
QY 9 DPAKCRYALGMQDRTIPDSISASS-----SWSDDTAARHSRLSSDGDGAWCPAGSV 61  
Db 1281 DLNCSMPLGMQNAISDSQITASSLSNIFATWSPSQ-----RLHLOGRTNNAWRPRVS- 1336  
QY 62 FPKEEYLQVLDLQRLHLVALVGTQGRHAGGLGKEFSRSLRYLRYSRDRGRRWGWKDRWGOE 121  
Db 1336 --SAEWMQLVDLQKTVKVTGTTGVKS--LLSSMYVKEFLVSSQDGRRTWTLFLQDGHK 1392  
QY 122 VISGNEDEPGVVLKDLGPPVAVLRVRYPRADRVMSVCLRVLYGCLWRD 171  
Db 1393 VFOGNDSTPTVNVNADPLPFLTRYLRHP-TSWAQHIALRLVLCGEAOD 1441

RESULT 9  
US-08-111-939-2  
; Sequence 2, Application US/08111939  
; Patent No. 5460951  
; GENERAL INFORMATION:  
; APPLICANT: Kawai, Shinji  
; APPLICANT: Takeshita, Sunao  
; APPLICANT: Okazaki, Makoto  
; APPLICANT: Amann, Egon  
; TITLE OF INVENTION: Bone-Related Carboxypeptidase-Like  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
; ADDRESSSEE: Dunner  
; STREET: 1300 I Street, N.W.  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005-3315  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/111,939  
; FILING DATE: 26-AUG-1993  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 324033/92  
; FILING DATE: 03-DEC-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 230029/92  
; FILING DATE: 28-AUG-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Forman, David S.  
; REGISTRATION NUMBER: 33,694  
; REFERENCE/DOCKET NUMBER: 02481.1321-00000  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-408-4000  
; TELEFAX: 202-408-4000  
; INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1128 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-111-939-2

Query Match 9.3%; Score 201.5; DB 1; Length 1128;  
Best Local Similarity 24.5%; Pred. No. 3.6e-12;  
Matches 80; Conservative 51; Mismatches 144; Indels 51; Gaps 13;

QY 12 KCRYALGMDRTIPDSISASSWSDSTAARHSRLSSDGDGAWCPAGSV 64  
DB 375 KCP-PIGMESHRIEDNQIRASSMRHGLGAQRGLNMQAGANEDDYDGAWCAE---S 430  
QY 65 EBEYLQVLDLQRLHLVALVGTQGRHAGLGKFEFSYRLRYSDGRWRWGWKDRWGQEVIS 124  
DB 431 QTMIEVDTRTRTFTGTGTQGRDS-STHDDFVTTFVFVGSNDSTQVMYMTNGYEEMTFY 489  
QY 125 GNEDEGVVLKDLGPPMVARLVREYPRADRVMSVCLRVLYGCLWRDGLLSYATPVGQTM 184  
DB 490 GNVKDTPLSELPEPVVARIRIPLTN-GSLCMRLVLC-----PVTVPV-YSY 539  
QY 185 YLSEAVYLNDSYDGTGTYGGLQYGLGQADGV-----VGLDDPRKSKQELRVW 232  
DB 540 YQNEVVTDSLDPRHH---SYKDMRQLMKAVNEECPTIIRTYSLG---KSSRGLKIY 591  
QY 233 -----PGYDYVGNWSNFSFGYVEMFEFDRLRAFOAQVHCNNMHTLGARLPGGV-E 284  
DB 592 AMEISDNPDGDELGPFRYTAGIHGNEV-LGRELLLLLMQYLQEQYRDGNPRVRLVQD 650  
QY 285 CRFRGRPAMWEGEPMRHNLGNLGD 310  
DB 651 TRIHLVPSLNDPGYEVAAQMGSEFEN 676

RESULT 10  
US-07-607-538C-5  
Sequence 5, Application US/07607538C  
Patent No. 5455031  
GENERAL INFORMATION:  
APPLICANT: Ceriani Dr., Roberto L.  
APPLICANT: Peterson Dr., Jerry A.  
APPLICANT: Larocca, David J.  
TITLE OF INVENTION: POLYPEPTIDE WITH 46  
TITLE OF INVENTION: DIFFERENTIATION ANTIGEN BINDING SPECIFICITY AND CLOTTING  
TITLE OF INVENTION: FACTORS V AND VIII LIGHT-CHAIN HOMOLOGIES,  
TITLE OF INVENTION: FUSION PROTEIN, POLYNUCLEOTIDE AND POLYRIBO-  
TITLE OF INVENTION: NUCLEOTIDE ENCODING THE POLYPEPTIDE, ANTI-  
TITLE OF INVENTION: POLYPEPTIDE ANTIBODIES, KITS AND METHODS OF  
TITLE OF INVENTION: USE THEREOF  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: V. Amzel & Assoc.  
STREET: 2055 No. 5455031th Broadway  
CITY: Walnut Creek  
STATE: California  
COUNTRY: USA  
ZIP: 94596  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS 5.0  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/607/538C  
FILING DATE: 01-NOV-1990  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Viviana Amzel  
REGISTRATION NUMBER: 30,930  
REFERENCE/DOCKET NUMBER: CRFCC-004

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (510) 943-1931  
TELEFAX: (510) 943-1189  
TELEX: N.A.  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 218 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE:  
US-07-607-538C-5

Query Match 8.9%; Score 193; DB 1; Length 218;

Best Local Similarity 31.8%; Pred. No. 2.6e-12;  
Matches 54; Conservative 33; Mismatches 67; Indels 16; Gaps 5;

QY 9 DPAKCRYALGMDRTIPDSISASS-----SWSDSTAARHSRLSSDGDGAWCPAGSV 61  
DB 56 DLNCSMPGLGMSKASDAQITASSYFTNMFATWSPSKARLHLOGRSN---AWRPQVN- 111  
QY 62 FKREYLVQVLDLQRLHLVALVGTQGRHAGLGKFEFSYRLRYSDGRWRWGWKDRWGQEV 121  
DB 111 --NPKEVLQVDFQKTMKVTGVTQGVKS--LLPTMYVKEFLISSQDGHQWTLFFQNGKVK 167  
QY 122 VLSGNEDEGVVLKDLGPPMVARLVREYPRADRVMSVCLRVLYGCLWRD 171  
DB 168 VFGNQDSFTPVVNSLDPLRLRYLRIHPQS-WVHQIALRMEVLGCEAQD 216

RESULT 11  
US-07-864-004B-4  
Sequence 4, Application US/07864004B  
Patent No. 5384771  
GENERAL INFORMATION:  
APPLICANT: Lollar, John S.  
APPLICANT: Runge, Marshall S.  
TITLE OF INVENTION: Hybrid Human/Porcine Factor VIII  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kilpatrick & Cody  
STREET: 1100 Peachtree Street  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: US  
ZIP: 30309  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/864,004B  
FILING DATE: 07 APRIL 1992  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: EMU106  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-815-6508  
TELEFAX: 404-815-6555  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2332 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO

```

Query Match      8.8%; Score 192; DB 1; Length 2332;
Best Local Similarity 31.8%; Pred. No. 9.8e-11;
Matches 54; Conservative 32; Mismatches 68; Indels 16; Gaps

QY      9 DPAKRYALGMQDRTPIDPSDISASS-----SWSNSTAARHSRLSESSDGDGAWCPAGSV 61
      |   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      2170 DLNCSMPLGMESKATSDAQITASSYFTNMFATWSPSKARLHIGRSN----AMRPQVN- 2225

QY      62 FPKREBYLQVDLQRHLVALVGTQGRHAGGLGKFEFSRYRLRYSRDGRNMWGWKDRNGQE 121
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      2225 --NPKEWLQVDFQKTMKVTGVTTCQVKS-LITSMYKVEFLITSSDQDGHOWTLFPONGKVK 2281

```



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 1999, 22:52:09 ; Search time 22.34 seconds  
(without alignments)  
423.042 Million cell updates/sec

Title: US-08-170-558-8  
Perfect score: 2175  
Sequence: 1 DADMGHDFPAKCRYALGMQ.....LELEPRGQVPVAKPEGSPTA 399

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database: A\_Geneseq\_36.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2167	99.6	919	1 R75502	Human mammary carc
2	2167	99.6	919	1 R75504	Human mammary carc
3	2167	99.6	919	1 W34672	Human mammary carc
4	2167	99.6	882	1 W34673	Human mammary carc
5	2167	99.6	876	1 W34675	Human mammary carc
6	2124.5	97.7	914	1 R71100	Protein-tyrosine-k
7	1002.5	46.1	855	1 R75503	Human colonic aden
8	1002.5	46.1	855	1 R75505	Human colonic aden
9	1002.5	46.1	855	1 W34671	CCR-2, a human mam
10	1002.5	46.1	855	1 W77114	Discoidin domain r
11	995.5	45.8	854	1 W79152	Receptor protein t
12	995.5	45.8	854	1 W81409	Receptor protein t
13	425	19.5	650	1 R54089	Partial sequence o
14	364	16.7	563	1 W34674	Human mammary carc
15	222	10.2	909	1 W96254	Mouse semaphorin r
16	222	10.2	926	1 W96255	Human semaphorin r
17	220	10.1	909	1 W96249	Rat semaphorin rec
18	220	10.1	909	1 W96250	Mouse semaphorin r
19	220	10.1	914	1 W96251	Mouse semaphorin r
20	220	10.1	926	1 W96252	Mouse semaphorin r
21	220	10.1	931	1 W96253	Mouse semaphorin r
22	220	10.1	901	1 W96256	Mouse semaphorin r
23	220	10.1	906	1 W96257	Mouse semaphorin r
24	217	10.0	925	1 W96308	Neuropilin-2, New
25	211.5	9.7	845	1 W36816	Human E2A-binding
26	210	9.7	1443	1 W44137	Homo sapiens facto
27	210	9.7	2133	1 W44133	Sus scrofa factor
28	204	9.4	923	1 W96246	Human semaphorin r
29	204	9.4	923	1 W96248	Mouse semaphorin r
30	201.5	9.3	1128	1 R49994	Mouse carboxypepti
31	201.5	9.3	1128	1 W36817	Mouse E2A-binding
32	199	9.1	921	1 W96247	Rat semaphorin rec
33	199	9.1	922	1 W96309	Neuropilin, New is
34	193	8.9	768	1 W36823	Human E2A-binding
35	193	8.9	2351	1 P70448	Human Factor VIII:
36	193	8.9	2351	1 R78223	Human Factor-VIII:
37	193	8.9	768	1 W36815	Human E2A-binding
38	192	8.8	1424	1 P91169	Sequence of 740 Ar
39	192	8.8	1424	1 P80268	Modified factor VI
40	192	8.8	1516	1 P80265	Modified factor VI
41	192	8.8	1425	1 P80267	Modified factor VI
42	192	8.8	2351	1 P81113	Factor VIII encode
43	192	8.8	2351	1 P80659	Sequence of human

#### ALIGNMENTS

RESULT 1

R75502  
ID R75502 standard; Protein: 919 AA.  
AC R75502;  
DT 26-NOV-1995 (first entry)  
DE Human mammary carcinoma kinase 10 (MCK-10).  
KW Mammary carcinoma kinase 10; transmembrane receptor;  
KW receptor tyrosine kinase; cancer.  
OS Homo sapiens.  
FH Key  
FT Location/Qualifiers  
FT peptide 1..18  
FT /label= signal  
FT domain 31..185  
FT /label= discoidin I-like domain  
FT cleavage\_site 304..307  
FT /label= putative precursor cleavage site  
FT region 417..439  
FT /label= transmembrane  
FT misc\_difference 505..541  
FT /label= alternatively spliced sequence I  
FT misc\_difference 666..671  
FT /label= alternatively spliced sequence II  
FT misc\_difference 25..42  
FT /label= NT alpha  
FT /note= "peptide antibody recognition site"  
FT misc\_difference 309..321  
FT /label= NT beta  
FT /note= "see above"  
FT misc\_difference 909..919  
FT /label= CT beta  
FT /note= "see above"

WO9514088-A.

26-MAY-1995.

16-NOV-1994; E03797.

16-NOV-1993; US-153397.

(PLAC) MAX PLANCK GES. FÖRDERUNG WISSENSCHAFTEN.

Alves FHE, Ullrich A;

WPI; 95-224054/29.

N-PSDB; Q92520.

New nucleic acid encoding MCK-10 receptor tyrosine kinase - and derived vectors, transformed cells, proteins and antibodies useful for diagnosis and treatment of proliferative disease, esp. cancer, and for screening modulators

Disclosure; Page 53-55; 115pp; English.

cDNA prep. from human breast cancer cell line MCF7 (ATCC HTB22) and used in a PCR with two degenerate oligo primer pools based on conserved sequences of the kinase domain of receptor tyrosine

kinases. One clone, designated MCK-10, was identified as novel RTK.

The PCR fragment was used to screen a lambda gtlI library of human fetal brain cDNA. Several overlapping clones were identified. The

composite of these cDNA clones is given in Q92520 and the deduced AA sequence in R75502. Some of the clones had a deletion of 6 AAs at

posn. 2315 in the MCK-10 sequence. MCK-10 has all the

characteristics of a receptor PTK (see R75502 FT). Screening of

human placental library yielded two cDNA clones MCK-10-1 and

MCK-10-2. One of the clones isolated from the human fetal brain

library contd. an additional 18 nts in the TK domain. The MCK-10 splice

isoforms have been designated MCK-10-1 (with an additional 111 bp between

nts 1832 and 1943); MCK-10-2 (without any insertions); and MCK-10-3 (with the

additional 111 bp and and 18 bp in the TK domain); and MCK-10-4 (with the

proreceptors are 101.13 and 97.17 kD respectively, and can thus be

subdivided into a 34.31 kD alpha subunit and and 66.84 or 62.88 kD

beta subunits that contain the TK homology and alternative splice sites.

Sequence 919 AA;

5Q

Factor VIII:c varia  
Human factor VIII

PT	New nucleic acid encoding CKK-2 receptor tyrosine kinase - and
PT	derived vectors, transformed cells, proteins and antibodies, useful
PT	for diagnosis and treatment of proliferative and nervous system
PS	diseases and for screening modulators
PS	Disclosure; Page 70-72; 115pp; English.
CC	cDNA prepd. from human breast cancer cell line MCF7 (ATCC HTB22) was
CC	used in a PCR with two degenerate oligo primer pools based on
CC	conserved sequences of the kinase domain of receptor tyrosine
CC	kinases. One clone, designated MCK-10, was identified as novel RTK.
CC	The PCR fragment was used to screen a lambda gt11 library of human
CC	fetal brain cDNA. Several overlapping clones were identified. The
CC	composite of these cDNA clones is given in Q92522 and the deduced AA
CC	sequence in R75504. Some of the clones had a deletion of 6AA at posn.
CC	2315 in the MCK-10 sequence. MCK-10 has all the characteristics of
CC	a receptor TK (see R75504 FT). Screening of human placental library
CC	yielded two cDNA clones. One of the clones isolated from the human
CC	fetal brain library contained an additional 18 nts in the TK
CC	domain. The MCK-10 splice isoforms have been designated MCK-10-1
CC	(without any insertions); MCK-10-3 (with the additional 111 bps and
CC	18 bp in the TK domain); and MCK-10-4 (with the additional 111 bps and
CC	the predicted mol. wts. of MCK-10-1 and MCK-10-2 predecessors are
CC	101.13 and 97.17 kD respectively, and can thus be subdivided into a
CC	34.31 kD alpha subunit and a 66.84 or 62.88 kD beta subunits that
CC	contain the TK homology and alternative splice sites.
CC	Sequence 919 AA;
SQ	

Query Match	99.8%; Score 2167; DB 1; Length 919;
Best Local Similarity	99.7%; Pred. No. 7.8e-210;
Matches 398; Conservative	0; Mismatches 1; Indels 0; Gaps 0;

QY	1	DADMGHFDPAKRYALGMDRTIPDSIDISASSWSDDTAARSLRLESSDGDGANCPAGS	60
DB			
DB	19	DADMGHFDPAKRYALGMDRTIPDSIDISASSWSDDTAARSLRLESSDGDGANCPAGS	78
QY	61	VFPKEBEYLQVDLQLRLHLVALVGTQGRHAGGLGKEFSRSYRLRAYSRDGRRWGMWKDRWGQ	120
DB			
DB	79	VFPKEBEYLQVDLQLRLHLVALVGTQGRHAGGLGKEFSRSYRLRAYSRDGRRWGMWKDRWGQ	138
QY	121	EVLISGNEDPBGVVLDLGPMPVARLVRFYPADRVMVSVCLRVLYGCLWRDGLLSLTAPV	180
DB			
DB	139	EVLISGNEDPBGVVLDLGPMPVARLVRFYPADRVMVSVCLRVLYGCLWRDGLLSLTAPV	198
QY	181	GOTMYLSEAYLNDSYTDGHTVGGLQYGGIQLADGVVGLDDFRKSOELRWPFYDYVGW	240
DB			
DB	199	GOTMYLSEAYLNDSYTDGHTVGGLQYGGIQLADGVVGLDDFRKSOELRWPFYDYVGW	258
QY	241	SNHSFSGYVEMEFEEDRLRAFOAMQVHCNNMTILGARLPGGVCECFRRRGPPANAWEGEPM	300
DB			
DB	259	SNHSFSGYVEMEFEEDRLRAFOAMQVHCNNMTILGARLPGGVCECFRRRGPPANAWEGEPM	318
QY	301	RHNLGNLGDPRARAVSVPLGGVRAREFLQCRFLFACGPLLLFSISFTSDVNNSSPALGG	360
DB			
DB	319	RHNLGNLGDPRARAVSVPLGGVRAREFLQCRFLFACGPLLLFSISFTSDVNNSSPALGG	378
QY	361	TTPPPAPWPFGPPPTNFSSLEPRGQQPVAKPEGSPGA	399
DB			
DB	379	TTPPPAPWPFGPPPTNFSSLEPRGQQPVAKPEGSPGA	417

RESULT	3
W34672	
ID	W34672 standard; Protein: 919 AA.
AC	W34672;
DT	17-FEB-1998 (first entry)
DE	Human mammary carcinoma kinase 10 (MCK-10) amino acid sequence.
KW	Mammary carcinoma kinase; MCK-10; receptor tyrosine kinase;
KW	proliferative disease; cancer; insulin receptor family;
KW	tyrosine kinase neurotrophin receptor; MCK-10 activity;
KW	neurological disorder; aberrant expression.
OS	Homo sapiens.
FH	Key Location/Qualifiers
Key	





FT /label= potential\_substrate\_binding\_site  
 FT 26..42  
 FT /note= "antibody recognition sequence NTalpha"  
 FT 309..321  
 FT /note= "antibody recognition sequence NTbeta"  
 FT 865..882  
 FT /note= "antibody recognition sequence CTbeta"  
 PN US5677144-A.  
 PD 14-OCT-1997.  
 PE 08-NOV-1994; 336343.  
 PR 16-NOV-1993; US-153397.  
 PA (ALVE/) ALVES F H E.  
 PA (ULLR/) ULLRICH A.  
 PI Alves FHE, Ullrich A;  
 DR WPI; 97-511869/47.  
 PT Truncated receptor tyrosine kinase CCK-2 - and nucleic acid coding  
 PT for it, useful for cancer diagnosis  
 PS Disclosure: Page -: 70pp; English.  
 CC The present sequence represents a splice variant of a mammary  
 CC carcinoma kinase (MCK-10). This kinase belongs to a novel family  
 CC of receptor tyrosine kinases, and expression is associated with  
 CC proliferative diseases such as cancer. The MCK-10 receptor tyrosine  
 CC kinase has extensive sequence similarity to the insulin receptor family.  
 CC The MCK-10 gene was isolated by PCR using 2 degenerate oligonucleotide  
 CC primer pools, using a template cDNA synthesised by reverse transcription  
 CC of poly-A RNA from the human mammary carcinoma cell line MCF7. The  
 CC amplified PCR product was used to screen human foetal brain and placental  
 CC libraries, from which the present splice variant was isolated. This  
 CC splice variant does not possess amino acids 505-541 of MCK-10 (W34672).  
 CC The sequence represented by amino acids 548-558 may be important, as  
 CC deletion of this motif in the activin receptor serine/threonine kinase  
 CC results in reduced ligand binding affinity. MCK-10 is expressed in brain  
 CC tissue, and the protein shares homology with the tyrosine kinase  
 CC neurotrophin receptor. Modulation of MCK-10 activity therefore may be used  
 CC for treatment of neurological disorders. MCK-10 is also expressed in a  
 CC variety of cancer cell lines and tumour tissue. The nucleotide sequence  
 CC of MCK-10, or parts of it, can be used for diagnostic purposes to detect  
 CC aberrant expression of MCK-10 genes. Inhibitors of MCK-10 (or splice  
 CC variants) receptor activity may have therapeutic value in the treatment  
 CC of diseases such as cancer.  
 CC note: the present sequence does not appear in the specification, but was  
 CC created using information provided.  
 SQ Sequence 882 AA;

Query Match 99.6%; Score 2167; DB 1; Length 882;  
 Best Local Similarity 99.7%; Pred. No. 7.4e-210;  
 Matches 398; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DADMKGFDPKCRYALGMDRTIPDSISASSWSUSTAAHRSRLSSDGDGAWCPAGS 60  
 DB 19 DADMKGFDPKCRYALGMDRTIPDSISASSWSUSTAAHRSRLSSDGDGAWCPAGS 78  
 QY 61 VFPKEEYLQVDLQRLHLVALVGTQGRHAGLGKFEFSRLYRSRDRGWRMKGDRWGQ 120  
 DB 79 VFPKEEYLQVDLQRLHLVALVGTQGRHAGLGKFEFSRLYRSRDRGWRMKGDRWGQ 138  
 QY 121 EYISGNEDEGVKLDGPPMVARLVRFYPRADRVMSVCLRVLYGCLWRDGLLSYTPV 180  
 DB 139 EYISGNEDEGVKLDGPPMVARLVRFYPRADRVMSVCLRVLYGCLWRDGLLSYTPV 198  
 QY 181 GOTMYLSEAVYLNDSYDGHVTGGLGGLQGLADGVVGLDDFRKSELRWPGDYVGV 240  
 DB 199 GOTMYLSEAVYLNDSYDGHVTGGLGGLQGLADGVVGLDDFRKSELRWPGDYVGV 258  
 QY 241 SNHFSFGYVEMEFDFRLRAFOAQMVHCNNHHTLGARLPGGVBCRRFRGPAMAWEGEPM 300  
 DB 259 SNHFSFGYVEMEFDFRLRAFOAQMVHCNNHHTLGARLPGGVBCRRFRGPAMAWEGEPM 318  
 QY 301 RHNLGNGLDPRARVSVPLGGRVARELQCRFLFAGPWLIFSEISFISDVYNNSSPALGG 360  
 DB 319 RHNLGNGLDPRARVSVPLGGRVARELQCRFLFAGPWLIFSEISFISDVYNNSSPALGG 378

QY 361 TFPAPWMPGPPPTNFSSLELEPRGQVPVAKPEGSPTA 399  
 DB 379 TFPAPWMPGPPPTNFSSLELEPRGQVPVAKPEGSPTA 417  
 RESULT 5  
 W34675  
 ID W34675 standard; Protein: 876 AA.  
 AC W34675;  
 DT 17-FEB-1998 (first entry)  
 DE Human mammary carcinoma kinase 10 (MCK-10) splice variant 3.  
 KW Mammary carcinoma kinase; MCK-10; receptor tyrosine kinase;  
 KW proliferative disease; cancer; insulin receptor family;  
 KW tyrosine kinase neurotrophin receptor; MCK-10 activity;  
 KW neurological disorder; aberrant expression.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Peptide 1..118  
 FT /label= signal\_sequence  
 FT 19..876  
 FT /note= "mature\_protein"  
 FT Domain 31..185  
 FT /label= Discoidin\_I\_like\_domain  
 FT Cleavage\_site 304..307  
 FT /label= endopeptidase\_furin  
 FT /note= "putative precursor cleavage site"  
 FT Region 48..439  
 FT /label= transmembrane\_region  
 FT Binding\_site 580..590  
 FT /label= ATP\_binding\_motif  
 FT Modified\_site 760..761  
 FT /label= autophosphorylation\_sites  
 FT /note= "putative"  
 FT Modified\_site 756..756  
 FT /label= autophosphorylation\_site  
 FT /note= "putative"  
 FT Binding\_site 802..805  
 FT /label= binding\_motif\_for\_P13\_kinase  
 FT /note= "binding motif for phosphatidylinositol 3'  
 FT kinase"  
 FT Binding\_site 790  
 FT /label= potential\_substrate\_binding\_site  
 FT 26..42  
 FT /note= "antibody recognition sequence NTalpha"  
 FT Region 309..321  
 FT /note= "antibody recognition sequence NTbeta"  
 FT Region 860..877  
 FT /note= "antibody recognition sequence CTbeta"  
 PN US5677144-A.  
 PD 14-OCT-1997.  
 PE 08-NOV-1994; 336343.  
 PR 16-NOV-1993; US-153397.  
 PA (ALVE/) ALVES F H E.  
 PA (ULLR/) ULLRICH A.  
 PI Alves FHE, Ullrich A;  
 DR WPI; 97-511869/47.  
 PT Truncated receptor tyrosine kinase CCK-2 - and nucleic acid coding  
 PT for it, useful for cancer diagnosis  
 PS Disclosure: Page -: 70pp; English.  
 CC The present sequence represents a splice variant of a mammary  
 CC carcinoma kinase (MCK-10). This kinase belongs to a novel family  
 CC of receptor tyrosine kinases, and expression is associated with  
 CC proliferative diseases such as cancer. The MCK-10 receptor tyrosine  
 CC kinase has extensive sequence similarity to the insulin receptor family.  
 CC The MCK-10 gene was isolated by PCR using 2 degenerate oligonucleotide  
 CC primer pools, using a template cDNA synthesised by reverse transcription  
 CC of poly-A RNA from the human mammary carcinoma cell line MCF7. The  
 CC amplified PCR product was used to screen human foetal brain and placental  
 CC libraries, from which the present splice variant was isolated. This  
 CC splice variant does not possess amino acids 505-541 or 666-671 of MCK-10  
 CC (W34672). The sequence represented by amino acids 548-558 may be  
 CC important, as deletion of this motif in the activin receptor  
 CC serine/threonine kinase results in reduced ligand binding affinity.

CC MCK-10 is expressed in brain tissue, and the protein shares homology with  
 CC the tyrosine kinase neurotrophin receptor. Modulation of MCK-10 activity  
 CC therefore may be used for treatment of neurological disorders. MCK-10 is  
 CC also expressed in a variety of cancer cell lines and tumour tissue. The  
 CC nucleotide sequence of MCK-10, or parts of it, can be used for diagnostic  
 CC purposes to detect aberrant expression of MCK-10 genes. Inhibitors of  
 CC MCK-10 (or splice variants) receptor activity may have therapeutic value  
 CC in the treatment of diseases such as cancer.  
 CC note: the present sequence does not appear in the specification, but was  
 CC created using information provided.  
 SQ Sequence 876 AA;

Query Match 99.6%; Score 2167; DB 1; Length 876;  
 Best Local Similarity 99.7%; Pred. No. 7.3e-210;  
 Matches 398; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 DADMKGHFDPAKCRYALGMQDRTPDSDISASSWSSTAAHRSRLSSDGDGAWCPAGS 60  
 DB 19 DADMKGHFDPAKCRYALGMQDRTPDSDISASSWSSTAAHRSRLSSDGDGAWCPAGS 78  
 QY 61 VFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKEFSRSYRLRYSRDGRMMGWKDRWGQ 120  
 DB 79 VFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKEFSRSYRLRYSRDGRMMGWKDRWGQ 138  
 QY 121 EVISGNEDEPGVVKDLGPPMVARLVRFYPRADRVMSVCLRVLYGCLWRDGLLSYTA 180  
 DB 139 EVISGNEDEPGVVKDLGPPMVARLVRFYPRADRVMSVCLRVLYGCLWRDGLLSYTA 198  
 QY 181 GQWYLSAYVLYNDSTYDGTGVLGGLQGLADGVVGLDDPRKSOELRWVPYDYGW 240  
 DB 199 GQWYLSAYVLYNDSTYDGTGVLGGLQGLADGVVGLDDPRKSOELRWVPYDYGW 258  
 QY 241 SNHSFSSGYVEMEFEDRLRAFOAMOVHCNNMHTLGARLPGGVCEFRFRGPAMAWEGEPM 300  
 DB 259 SNHSFSSGYVEMEFEDRLRAFOAMOVHCNNMHTLGARLPGGVCEFRFRGPAMAWEGEPM 318  
 QY 301 RHNGLGNLGDPRARAVSVPLGGVAREFLQCRFLFAGPWLIFSEISFISDVVNSSPALGG 360  
 DB 319 RHNGLGNLGDPRARAVSVPLGGVAREFLQCRFLFAGPWLIFSEISFISDVVNSSPALGG 378  
 QY 361 TFPAPWPPGPPPTNFSSLELEPRGQPVAKPEGSPTA 399  
 DB 379 TFPAPWPPGPPPTNFSSLELEPRGQPVAKPEGSPTA 417

RESULT 6  
 R71100  
 ID R71100 standard; Protein; 914 AA.  
 AC R71100;  
 DT 17-AUG-1995 (first entry)  
 DE Protein-tyrosine-kinase PTK22.  
 KW Protein-tyrosine-kinase; PTK; discoidin domain receptor; cancer;  
 KW breast tumor; mamma carcinoma; diagnosis; prognosis; therapy.  
 OS Homo sapiens.  
 PN WO9502187-A.  
 PD 19-JAN-1995.  
 PF 08-JUL-1994; G01480.  
 PR 09-JUL-1993; GB-014271.  
 PA (CANC-) CANCER RES INST.  
 PA (WELL) WELLCOME FOUND-LTD.  
 PI Barker KT, Crompton MR, Gusterson BA, Martindale JE;  
 PI Mitchell PJ, Page MJ, Spence P;  
 DR WPI: 95-066991/09.  
 DR N-PSDB: Q84782.  
 PT Method for screening substances, using protein tyrosine kinase -  
 PT for potential utility as therapeutic agents for cancer  
 PS Disclosure; Page 26-30; 51pp; English.  
 CC cDNA derived from tumor metastatic tissue was amplified using  
 CC primers (given in Q84783-84) based on sequences (R71101, R71103)  
 CC associated with protein-tyrosine-kinases (PTK). Novel PTK22 was  
 CC identified in an isolated subclone. The 3' sequence of PTK22 was  
 CC obtained by reverse transcription (using the primer of Q84786) and

CC PCR amplification (primers Q84787-88) of RNA of human breast  
 CC carcinoma cell line MDA MG 468. The partial DNA sequence of PTK22  
 CC is given in Q84782.  
 SQ Sequence 914 AA;

Query Match 97.7%; Score 2124.5; DB 1; Length 914;  
 Best Local Similarity 98.2%; Pred. No. 1.5e-205;  
 Matches 393; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 1 DADMKGHFDPAKCRYALGMQDRTPDSDISASSWSSTAAHRSRLSSDGDGAWCPAGS 60  
 DB 19 DADMKGHFDPAKCRYALGMQDRTPDSDISASSWSSTAAHRSRLSSDGDGAWCPAGS 78  
 QY 61 VFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKEFSRSYRLRYSRDGRMMGWKDRWGQ 120  
 DB 79 VFPKEEYLOVDLQRLHLVALVGTQGRHAGGLGKEFSRSYRLRYSRDGRMMGWKDRWGQ 138  
 QY 121 EVISGNEDEPGVVKDLGPPMVARLVRFYPRADRVMSVCLRVLYGCLWRDGLLSYTA 180  
 DB 139 EVISGNEDEPGVVKDLGPPMVARLVRFYPRADRVMSVCLRVLYGCLWRDGLLSYTA 198  
 QY 181 GQWYLSAYVLYNDSTYDGTGVLGGLQGLADGVVGLDDPRKSOELRWVPYDYGW 240  
 DB 199 GQWYLSAYVLYNDSTYDGTGVLGGLQGLADGVVGLDDPRKSOELRWVPYDYGW 258  
 QY 241 SNHSFSSGYVEMEFEDRLRAFOAMOVHCNNMHTLGARLPGGVCEFRFRGPAMAWEGEPM 300  
 DB 259 SNHSFSSGYVEMEFEDRLRAFOAMOVHCNNMHTLGARLPGGVCEFRFRGPAMAWEGEPM 318  
 QY 301 RHNGLGNLGDPRARAVSVPLGGVAREFLQCRFLFAGPWLIFSEISFISDVVNSSPALGG 360  
 DB 319 RHNGLGNLGDPRARAVSVPLGGVAREFLQCRFLFAGPWLIFSEISFISDVVNSSPALGG 378  
 QY 361 TFPAPWPPGPPPTNFSSLELEPRGQPVAKPEGSPTA 399  
 DB 379 TFPAPWPPGPPPTNFSSLELEPRGQPVAKPEGSPTA 418

RESULT 7  
 R75503  
 ID R75503 standard; Protein; 855 AA.  
 AC R75503;  
 DT 26-NOV-1995 (first entry)  
 DE Human colonic adenocarcinoma kinase 2 (CCK-2).  
 KW Mammary carcinoma kinase 10; MCK-10; transmembrane receptor; CCK-2;  
 KW receptor tyrosine kinase; colonic adenocarcinoma kinase 2; cancer.  
 OS Homo sapiens.  
 PN WO9514088-A.  
 PD 26-MAY-1995.  
 PF 16-NOV-1994; E03797.  
 PR 16-NOV-1993; US-153397.  
 PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.  
 PI Alves FHE, Ullrich A;  
 DR WPI: 95-224054/29.  
 DR N-PSDB: Q92521.  
 PT New nucleic acid encoding MCK-10 receptor tyrosine kinase - and  
 PT derived vectors, transformed cells, proteins and antibodies useful  
 PT for diagnosis and treatment of proliferative disease, esp. cancer,  
 PT and for screening modulators  
 PS Disclosure; Page 57-60; 115pp; English.  
 CC A member of the mammary carcinoma kinase 10 (MCK-10) receptor  
 CC tyrosine kinase family was identified using a PCR (with two degenerate  
 CC oligo primer pools based on conserved sequences of the kinase domain of  
 CC receptor tyrosine kinases) and cDNA prep. from colonic  
 CC adenocarcinoma RNA. The nt sequence of the novel receptor,  
 CC designated CCK-2, is given in Q92521 and the deduced AA sequence in  
 CC R75503. Analysis of the CCK-2 nt and AA sequence indicated  
 CC significant homology with MCK-10 throughout the extracellular,  
 CC transmembrane and intracellular regions. The regions of homology  
 CC extend into the N-terminus consensus sequence for the discoidin I  
 CC like family of proteins.  
 SQ Sequence 855 AA;

Query Match		46.1%;	Score 1002.5;	DB 1;	Length 855;
Best Local Similarity		53.6%;	Pred. No. 1.8e-92;		
Matches 196;		Conservative 54;	Mismatches 109;	Indels 7;	Gaps 6;
QY	5	KGHFDPAKRYALGMDRTIPDSISASSWSDDTAARHSRLSSDGDGAWCPAGSVFPK	64		
DB	22	KAQVNPACRYPLGMSGGQIPDEDITASSQWSESTAAYKGRLDSEGDGAWCPETPVPD	81		
QY	65	E-EYLVQDLQRLHLVALVGTQGRHAGLGKGFESRSYRLYSRDRGRWGWKDRWGQEV	123		
DB	82	DUKEFLQIDLHLTHLFTITLVGTQGRHAGHGIEFAPMYKINYSROGTRWISNRNRHGQVL	141		
QY	124	SGNEDEPGVWLKDLGPPMVARLVFYPADRVMSVCLRVLYGLWRDGLLSYAPVGQT	183		
DB	142	DGNSNPYDIFLKDLEPPIVAREVRFIPYTDHSMNVCMRVLYGCVWLDGLVSNAPAGQ	201		
QY	184	MYL--SEAVYLNDSYDGHVTGGYGLGQLADGVGLDFFRKSQBLRVWPGYDYVGS	241		
DB	202	FVLPGGSIYLNDSYDVG-AVGYSMTGLGQLTDGVSGLDDFTQTHEYHVWPGYDYVGR	260		
QY	242	NHSFSSGYVEMEFEDRLRAFOAMQVHCNNMHTLGARLPGGVCEFRFRGPAMAWEGEPMR	301		
DB	261	NESATNGYIEIMFEFDRIRNFTTMKVHCNNMFAKGVKIFKEVQCYF-RSEASEWEPNAIS	319		
QY	302	HNLGNLGDPRARAVSVPLGGRVARFLQCFRFLAGPWLFLSEISFISD-VYNNSSPALGG	360		
DB	320	FPLVLDVNPASRFVTVPLHHRMASAKCOYHFADTWMFSEITFQSDAAMYNSEAL-P	378		
QY	361	TFPPAP 366			
DB	379	TSPMAP 384			
RESULT 9					
ID	W34671	standard; Protein: 855 AA.			
AC	W34671;				
DT	16-FEB-1998	(first entry)			
DE	CKK-2, a human mammary carcinoma kinase 10 MCK-10) family member,				
KW	Mammary carcinoma kinase; MCK-10; CKK-2; receptor tyrosine kinase;				
KW	Proliferative disease; cancer; MCK-10 activity; aberrant expression.				
OS	Homo sapiens.				
FH	Key	Location/Qualifiers			
FT	Modified_site	121	/note= "N-glycosylated"		
FT	Modified_site	213	/note= "N-glycosylated"		
FT	Modified_site	261	/note= "N-glycosylated"		
FT	Modified_site	280	/note= "N-glycosylated"		
FT	Modified_site	328	/note= "N-glycosylated"		
FT	Modified_site	372	/note= "N-glycosylated"		
FT	Modified_site	503	/note= "N-glycosylated"		
FT	Modified_site	736	/note= "putative autophosphorylation and substrate binding site"		
FT	Modified_site	740	/note= "putative autophosphorylation and substrate binding site"		
FT	Modified_site	741	/note= "putative autophosphorylation and substrate binding site"		
FT	Modified_site	813	/note= "putative autophosphorylation and substrate binding site"		
FT	Modified_site	825	/note= "putative autophosphorylation and substrate binding site"		
FT	Modified_site		/note= "putative autophosphorylation and substrate binding site"		

Query Match		46.1%;	Score 1002.5;	DB 1;	Length 855;
Best Local Similarity		53.6%;	Pred. No. 1.8e-92;		
Matches 196;		Conservative 54;	Mismatches 109;	Indels 7;	Gaps 6;
QY	5	KGHFDPAKRYALGMDRTIPDSISASSWSDDTAARHSRLSSDGDGAWCPAGSVFPK	64		
DB	22	KAQVNPACRYPLGMSGGQIPDEDITASSQWSESTAAYKGRLDSEGDGAWCPETPVPD	81		
QY	65	E-EYLVQDLQRLHLVALVGTQGRHAGLGKGFESRSYRLYSRDRGRWGWKDRWGQEV	123		
DB	82	DUKEFLQIDLHLTHLFTITLVGTQGRHAGHGIEFAPMYKINYSROGTRWISNRNRHGQVL	141		
QY	124	SGNEDEPGVWLKDLGPPMVARLVFYPADRVMSVCLRVLYGLWRDGLLSYAPVGQT	183		
DB	142	DGNSNPYDIFLKDLEPPIVAREVRFIPYTDHSMNVCMRVLYGCVWLDGLVSNAPAGQ	201		
QY	184	MYL--SEAVYLNDSYDGHVTGGYGLGQLADGVGLDFFRKSQBLRVWPGYDYVGS	241		
DB	202	FVLPGGSIYLNDSYDVG-AVGYSMTGLGQLTDGVSGLDDFTQTHEYHVWPGYDYVGR	260		
QY	242	NHSFSSGYVEMEFEDRLRAFOAMQVHCNNMHTLGARLPGGVCEFRFRGPAMAWEGEPMR	301		
DB	261	NESATNGYIEIMFEFDRIRNFTTMKVHCNNMFAKGVKIFKEVQCYF-RSEASEWEPNAIS	319		
QY	302	HNLGNLGDPRARAVSVPLGGRVARFLQCFRFLAGPWLFLSEISFISD-VYNNSSPALGG	360		
DB	320	FPLVLDVNPASRFVTVPLHHRMASAKCOYHFADTWMFSEITFQSDAAMYNSEAL-P	378		
QY	361	TFPPAP 366			
DB	379	TSPMAP 384			
RESULT 8					
ID	R75505	standard; Protein: 855 AA.			
AC	R75505;				
DT	26-NOV-1995	(first entry)			
DE	Human colonic adenocarcinoma kinase 2 (CKK-2).				
KW	Mammary carcinoma kinase 10; MCK-10; transmembrane receptor; CKK-2;				
KW	receptor tyrosine kinase; colonic adenocarcinoma kinase 2.				
OS	Homo sapiens.				
PN	W09514089-A.				
PD	26-MAY-1995.				
PF	16-NOV-1994; E03799.				
PR	16-NOV-1993; US-153397.				
PA	(PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.				
PI	Alves FHE, Ullrich A;				
DR	WPI; 95-224055/29.				
DR	N-PSDB; Q92523.				
PT	New nucleic acid encoding CKK-2 receptor tyrosine kinase - and				
PT	derived vectors, transformed cells, proteins and antibodies, useful				
PT	for diagnosis and treatment of proliferative and nervous system				
PT	diseases and for screening modulators				
PS	Disclosure: Page 74-77; 115pp; English.				
CC	A member of the mammary carcinoma kinase 10 (MCK-10) receptor				
CC	tyrosine kinase family was identified using a PCR (with two				
CC	degenerate oligo primer pools based on conserved sequences of the				
CC	kinase domains of receptor tyrosine kinases) and cDNA prep. from				
CC	colonic adenocarcinoma RNA. The nt sequence of the novel receptor,				
CC	designated CKK-2, is given in Q92523 and the deduced AA sequence in				
CC	R75505. Analysis of CKK-2 nt and AA sequences indicated significant				
CC	homology with MCK-10 throughout the extracellular, transmembrane				
CC	and intracellular regions. The regions of homology extend into the				
CC	N-terminus consensus sequence of the discoidin I like family of				
CC	proteins. CKK-2 was predominantly found in all stromal cells				
CC	whereas MCK-10 expression was strongly confined to neoplastic				
CC	cells themselves. Between the two RTKs, the juxtamembrane region				
CC	is the region of most extensive sequence divergence.				
CC	Sequence 855 AA;				
SQ					

FT Region 400..421 binding site"  
FT /label= transmembrane\_region  
FT /note= "putative"  
FT 30..185  
FT /label= Discoidin\_I\_like\_domain  
FT 433..438  
FT /label= protein\_kinase\_C\_binding\_site  
FT /note= "putative"  
PN US5677144-A.  
PD 14-OCT-1997.  
PF 08-NOV-1994; 336343.  
PR 16-NOV-1993; US-153397.  
PA (ALVE/) ALVES F H E.  
PA (ULLR/) ULLRICH A.  
PI Alves FHE, Ullrich A;  
DR WPI; 97-511869/47.  
DR N-PSDB; T93784.  
PT Truncated receptor tyrosine kinase CCK-2 - and nucleic acid coding  
PT for it, useful for cancer diagnosis  
PS Claim 5; Fig 3; 70pp; English.  
CC The present sequence represents the amino acid sequence of human CCK-2, a  
CC member of the mammary carcinoma kinase 10 (MCK-10, W34672) family of  
CC receptor tyrosine kinases. The protein contains a remarkably high  
CC number of proline residues arranged as PXXP or PXXR repeats, suggesting  
CC a random coil structure for the hydrophilic juxtamembrane region. This  
CC region is probably a major domain for interactions with cellular  
CC substrates and other regulatory proteins. Expression of CCK-2 is  
CC associated with proliferative diseases such as cancer. The CCK-2 gene was  
CC identified by PCR and a cDNA prepared from colonic adenocarcinoma RNA.  
CC CCK-2 is expressed in a wide variety of cancer cell lines and tumour  
CC tissue. The CCK-2 nucleic acids can be used for diagnostic purposes to  
CC detect aberrant expression of CCK-2 genes. Engineered cell lines,  
CC containing recombinant vectors with the present sequence, are useful for  
CC producing infectious retroviral particles. The cell lines may also be  
CC used to evaluate and screen drugs involved in CCK-2 activation and  
CC regulation.  
SQ Sequence 855 AA;

Query Match 46.1%; Score 1002.5; DB 1; Length 855;  
Best Local Similarity 53.6%; Pred. No. 1.8e-92;  
Matches 196; Conservative 54; Mismatches 109; Indels 7; Gaps 6;  
QY 5 KGHDPKACRYALGMDRTIPDSISASSWSSTAAHRSRLESSDGDGAWCPAGSVFPK 64  
DB 22 KAQVNPALCRPLGMSGGQIPEDITASSQWSESTAKYGRDSEGDGAWCPPIPEPD 81  
QY 65 E-EEYLDVQLRHLVALVGTQGRHAGLGKGFERSYRLRYSRDGRMMGWKDRWGQEV 123  
DB 82 DLKEFLQIDLHTLFTLVGTQGRHAGHGIEFAPWKINYSRDGTWISWRNRHGKQVL 141  
QY 124 SGNEDPGVVLKDLGPPMVARLRFYPRADRVMSVCLRVLYGCLWRDGLLSYTPVGT 183  
DB 142 DGSNPYDIFLKDLEPPIVARFVRFPVTDHSMNVMCRVLYGCVLWDLGLVSYNAPAQ 201  
QY 184 MYL--SEAVYLDSTYDGTGGVGLQYGLGOLADGVVGLDFRKSQELRVWPGDYVGS 241  
DB 202 FVLPGGSIYLDNSYDG-AVGYSWTEGLGQLTGVSGLDLDTOTHEHYHWPGDYVGNR 260  
QY 242 NHSSSGYVMEFEFDRLAFQAMQVHCNNHTIGARLPGGVECFRRFGPAMAWEGEPMR 301  
DB 261 NESATNGYIEIMFEDRIRNFTTKVHCNMFAGVKIFKEVQCYF-RSEASEWEPNAIS 319  
QY 302 HNLGNLGDPRARAVSVPLGGRVARFLOCRFLFAGPWLLESEISFTSD-VYNNSSPALGG 360  
DB 320 FPLVLDVNPASAREVTVPVPLHHRMASAIKCOYHFDATWMMFSEITFQSDAAMYNSEAL-P 378  
QY 361 TFPAP 366  
DB 379 TSPMAP 384

RESULT 11

W79152  
ID W79152 standard; Protein; 854 AA.  
AC W79152;  
DT 19-NOV-1998 (first entry)  
DE Receptor protein tyrosine kinase (PTK) subtype tyro-10.  
KW PTK; receptor; protein tyrosine kinase; brain tissue.

RESULT 10  
W77114  
ID W77114 standard; Protein; 855 AA.  
AC W77114;  
DT 16-NOV-1998 (first entry)  
DE Discoidin domain receptor 2 protein.  
KW Discoidin domain receptor; transformation; metastasis; collagen; ss;  
KW Cleidocranial dysplasia; Sicker syndrome; extracellular matrix; MMP-1.  
OS Homo sapiens.  
PN W09834954-A2.  
PD 13-AUG-1998.  
PF 05-FEB-1998; CA0093.  
PR 06-FEB-1997; US-041578.  
PA (MOUN) MOUNT SINAI HOSPITAL CORP.  
PI Pawson A, Vogel W;  
DR WPI; 98-447168/38.  
DR N-PSDB; V48292.  
PT Novel ligands of discoidin domain receptor tyrosine kinase,  
PT especially collagen - useful for treating e.g. metastasis,  
PT cleidocranial dysplasia or Sicker syndrome  
PS Disclosure: Fig 22a; 115pp; English.  
CC The discoidin domain receptor (DDR) can be used to identify and evaluate  
CC substances which affect DDR receptor tyrosine kinase signalling pathways  
CC in the cell. Compounds which modulate such signalling pathways can be  
CC used to alter transformation or metastasis in mammals, to treat  
CC conditions involving structural or functional deregulation of collagens,  
CC e.g. Cleidocranial dysplasia or Sicker syndrome, conditions requiring  
CC modulation of extracellular matrix synthesis, degradation or remodelling,  
CC or to treat conditions needing modulation of MMP-1 expression such as  
CC wound healing.  
SQ Sequence 855 AA;

Query Match 46.1%; Score 1002.5; DB 1; Length 855;  
Best Local Similarity 53.6%; Pred. No. 1.8e-92;  
Matches 196; Conservative 54; Mismatches 109; Indels 7; Gaps 6;

QY 5 KGHDPKACRYALGMDRTIPDSISASSWSSTAAHRSRLESSDGDGAWCPAGSVFPK 64  
DB 22 KAQVNPALCRPLGMSGGQIPEDITASSQWSESTAKYGRDSEGDGAWCPPIPEPD 81  
QY 65 E-EEYLDVQLRHLVALVGTQGRHAGLGKGFERSYRLRYSRDGRMMGWKDRWGQEV 123  
DB 82 DLKEFLQIDLHTLFTLVGTQGRHAGHGIEFAPWKINYSRDGTWISWRNRHGKQVL 141  
QY 124 SGNEDPGVVLKDLGPPMVARLRFYPRADRVMSVCLRVLYGCLWRDGLLSYTPVGT 183  
DB 142 DGSNPYDIFLKDLEPPIVARFVRFPVTDHSMNVMCRVLYGCVLWDLGLVSYNAPAQ 201  
QY 184 MYL--SEAVYLDSTYDGTGGVGLQYGLGOLADGVVGLDFRKSQELRVWPGDYVGS 241  
DB 202 FVLPGGSIYLDNSYDG-AVGYSWTEGLGQLTGVSGLDLDTOTHEHYHWPGDYVGNR 260  
QY 242 NHSSSGYVMEFEFDRLAFQAMQVHCNNHTIGARLPGGVECFRRFGPAMAWEGEPMR 301  
DB 261 NESATNGYIEIMFEDRIRNFTTKVHCNMFAGVKIFKEVQCYF-RSEASEWEPNAIS 319  
QY 302 HNLGNLGDPRARAVSVPLGGRVARFLOCRFLFAGPWLLESEISFTSD-VYNNSSPALGG 360  
DB 320 FPLVLDVNPASAREVTVPVPLHHRMASAIKCOYHFDATWMMFSEITFQSDAAMYNSEAL-P 378  
QY 361 TFPAP 366  
DB 379 TSPMAP 384

RESULT 11

W79152  
ID W79152 standard; Protein; 854 AA.  
AC W79152;  
DT 19-NOV-1998 (first entry)  
DE Receptor protein tyrosine kinase (PTK) subtype tyro-10.  
KW PTK; receptor; protein tyrosine kinase; brain tissue.



```

FT modified_site 167 /note= "Potential N-linked glycosylation site."
FT domain 195. .216 /label= Transmembrane domain.
FT region 365. .370 /label= ATP binding region.
FT /note= "Highly conserved among protein tyrosine
FT kinase enzymes."
FT
PN DE4239817-A.
PD 01-JUN-1994.
PD 26-NOV-1992; 239817.
PF 26-NOV-1992; DE-239817.
PR (CHEM-) CHEMOTHERAPEUTISCHES FORSCHUNG.
PI Holtrich U, Ruebsamen-waigmann H, Strebhardt K;
DR WPI; 94-184380/23.
DR N-PSDB; 064158.
PT New protein tyrosine kinase and related nucleic acid - vectors,
PT transformed cells, etc., useful for diagnosis and treatment of
PT tumours
PS Claim 1; Page 8; 9pp; German.
CC The gene encoding this polypeptide is related to the trk
CC proto-oncogene. Antibodies against the polypeptide are useful for
CC diagnosis and for the treatment of tumours. The antibodies may also
CC be radiolabelled or coupled to a cytotoxin for destruction of cancer
CC cells. Antisense nucleic acid can be used to inhibit gene
CC expression.
CC Sequence 650 AA;
SQ
Query Match 19.5%; Score 425; DB 1; Length 650;
Best Local Similarity 50.6%; Pred. No. 2e-34;
Matches 90; Conservative 26; Mismatches 58; Indels 4; Gaps 4;
QY 190 VYLNDSYDGHVGGYGLQGLADGVGLDDFRKSOELRWPGYDYVGNHFSFGY 249
DB :||||| ||| || ||||| ||| ||||| ::| ||||| ||| ||| :||
5 IYLNDSYDGHVGGYGLQGLADGVGLDDFTOTHEYVWPGYDYVGNHFSFGY 63
QY 250 VEEFFEDRLRAQAMQVHCHNNMHTLGARLPGGVCEFRGRGPMAMWEGPMMHNLG 309
DB ::||| ||| :| ||||| ||| ::| ||| ||| ||| :|
64 IEIMFEDRLRNTTKMVCNNMFAKGVKFEVQCF-RSEASEWPAISFFPLVDV 122
QY 310 DPRARVSVPLGGRVAFRLQCRFLRAGPMLLSEISFISD-VYNNSSPALGGTFFPAP 366
DB :||| ||| ||| :||| ||| ||| ||| ||| ||| ||| ||| |||
123 NPSARFVTVPLHHRMASAIKQYHFADTWMMFSEITFQSDAAMYNSEAL-PTSPMAP 179
RESULT 14
W34674
ID W34674 standard; Protein; 563 AA.
AC W34674;
DT 17-FEB-1998 (first entry)
DE Human mammary carcinoma kinase 10 (MCK-10) splice variant 2..
KW Mammary carcinoma kinase; MCK-10; receptor tyrosine kinase;
KW proliferative disease; cancer; insulin receptor family;
KW tyrosine kinase neurotrophin receptor; MCK-10 activity;
KW neurological disorder; aberrant expression.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 1. .18
FT /label= signal_sequence
FT Protein 19. .919
FT /note= "mature_protein"
FT Domain 31. .185
FT /label= Discoidin_I_like_domain
FT Cleavage_site 304. .307
FT /label= endopeptidase_furin
FT /note= "putative precursor cleavage site"
FT Region 48. .439
FT /label= transmembrane_region
FT Binding_site 617. .627
FT /label= ATP_binding_motif
FT Modified_site 797. .798
FT /label= autophosphorylation_sites

```

```

FT Modified_site 793 /note= "putative"
FT /label= autophosphorylation_site
FT /note= "putative"
FT Binding_site 839. .842
FT /label= binding_motif_for_P13_kinase
FT /note= "binding motif for phosphatidylinositol 3'
FT kinase"
FT Binding_site 827. .827
FT /label= potential_substrate_binding_site
FT Binding_site 506. .509
FT /label= putative_receptor_binding_site_for_SHC
FT /note= "SHC is an oncogenic SH2 domain containing
FT molecule"
FT Binding_site 510. .513
FT /label= GTPase_activity_protein_binding_site
FT /note= "putative"
FT Region 26. .42
FT /note= "antibody recognition sequence NTalpha"
FT Region 309. .321
FT /note= "antibody recognition sequence NTbeta"
FT Region 897. .913
FT /note= "antibody recognition sequence CTbeta"
FT US5677144-A.
PN 14-OCT-1997.
PD 08-NOV-1994; 336343.
PF 16-NOV-1993; US-153397.
PR (ALVE/) ALVES F H E.
PA (ULLR/) ULLRICH A.
PI Alves FHE, Ullrich A;
DR WPI; 97-511869/47.
PT Truncated receptor tyrosine kinase CCK-2 - and nucleic acid coding
PT for it, useful for cancer diagnosis.
PT Disclosure; Page -: 70pp; English.
CC The present sequence represents a splice variant of a mammary
CC carcinoma kinase (MCK-10). This kinase belongs to a novel family
CC of receptor tyrosine kinases, and expression is associated with
CC proliferative diseases such as cancer. The MCK-10 receptor tyrosine
CC kinase has extensive sequence similarity to the insulin receptor family.
CC The MCK-10 gene was isolated by PCR using 2 degenerate oligonucleotide
CC primer pools, using a template cDNA synthesised by reverse transcription
CC of poly-A RNA from the human mammary carcinoma cell line MCF7. The
CC amplified PCR product was used to screen human foetal brain and placental
CC libraries, from which the present splice variant was isolated. This
CC splice variant does not possess amino acids 666-671 of MCK-10 (W34672).
CC The sequence represented by amino acids 585-595 may be important, as
CC deletion of this motif in the activin receptor serine/threonine kinase
CC results in reduced ligand binding affinity. MCK-10 is expressed in brain
CC tissue, and the protein shares homology with the tyrosine kinase
CC neurotrophin receptor. Modulation of MCK-10 activity therefore may be used
CC for treatment of neurological disorders. MCK-10 is also expressed in a
CC variety of cancer cell lines and tumour tissue. The nucleotide sequence
CC of MCK-10, or parts of it, can be used for diagnostic purposes to detect
CC aberrant expression of MCK-10 genes. Inhibitors of MCK-10 (or splice
CC variants) receptor activity may have therapeutic value in the treatment
CC of diseases such as cancer.
CC note: the present sequence does not appear in the specification, but was
CC created using information provided.
SQ Sequence 563 AA;
Query Match 16.7%; Score 364; DB 1; Length 563;
Best Local Similarity 98.5%; Pred. No. 2.4e-28;
Matches 66; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 333 LFAGPWLFFSEISFISDVNNSSPALGGTFFPAPWPPGPPPTNFSSLELEPRGQPVAK 392
DB :||||| ||| ||||| ||| ||||| ||| ||||| ||| ||||| |||
1 LFAGPWLFFSEISFISDVNNSSPALGGTFFPAPWPPGPPPTNFSSLELEPRGQPVAK 60
QY 393 PEGSPTA 399
DB :|||||
61 AEGSPTA 67

```





GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 4, 1999, 05:43:21 ; Search time 1266.11 seconds  
(without alignments)  
9135.687 Million cell updates/sec

Title: US-08-170-558-3  
Perfect score: 3637  
Sequence: 1 GAATTCGAGTCGACGTTG.....AAAAA..... 3637

Scoring table: IDENTITY\_NUC

Searched: 679419 seqs, 1590154680 residues

Database : GenEmbl.\*

- 1: gb\_ba1.\*
- 2: gb\_ba2.\*
- 3: gb\_om.\*
- 4: gb\_ov.\*
- 5: gb\_pat.\*
- 6: gb\_ph.\*
- 7: gb\_pl1.\*
- 8: gb\_pl2.\*
- 9: gb\_pr1.\*
- 10: gb\_pr2.\*
- 11: gb\_pr3.\*
- 12: gb\_ro.\*
- 13: gb\_st.\*
- 14: gb\_sts.\*
- 15: gb\_sy.\*
- 16: gb\_un.\*
- 17: gb\_vi.\*
- 18: em\_fun.\*
- 19: em\_htg.\*
- 20: em\_hum1.\*
- 21: em\_hum2.\*
- 22: em\_in.\*
- 23: em\_om.\*
- 24: em\_or.\*
- 25: em\_ov.\*
- 26: em\_pat.\*
- 27: em\_ph.\*
- 28: em\_pl.\*
- 29: em\_ro.\*
- 30: em\_sts.\*
- 31: em\_sy.\*
- 32: em\_un.\*
- 33: em\_vi.\*
- 34: gb\_htg1.\*
- 35: gb\_htg2.\*
- 36: gb\_in1.\*
- 37: gb\_in2.\*
- 38: gb\_ba1.\*
- 39: gb\_ba2.\*
- 40: em\_hum3.\*
- 41: em\_hum4.\*
- 42: gb\_pr4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match Length	DB ID	Description
1	3637	100.0	3637 5	I80845 Sequence 3

2	3451	94.9	3962	5	I68738
3	3441.8	94.6	3754	5	A42378
4	3438.6	94.5	3738	9	HUMRTK
5	3399.6	93.5	3803	9	HUMCAK
6	3243.6	89.2	3554	10	HSTRKE
7	3232.4	88.9	3841	10	HSRTYK1
8	3167.6	87.1	3692	10	HUMCAK
9	2375.8	65.3	3659	12	AF026259
10	2354.8	64.7	3674	12	MUSCAK
11	2314	63.6	3743	12	RATPTK3D
12	1197	32.9	1197	5	I80847
c	837.2	23.0	48731	11	AC004211
13	826.6	22.7	1072	10	HSRTK17
14	803.6	22.1	12010	11	HSU48705
15	669.4	18.4	3259	12	MTYR010
16	669.4	18.4	3259	12	MTYR010
17	666.8	18.3	3157	5	I68739
c	666.8	18.3	3157	5	I68740
18	666.8	18.3	3096	10	HSRPTK
19	570.4	15.7	660	12	S77585
20	340	9.3	549	12	S77586
21	326.6	9.0	845	10	HSRPTK6
22	250.2	6.9	970	10	HSRPTK12
23	250.2	6.9	970	10	HSRPTK12
c	250	6.7	355	10	G15457
24	243	6.7	355	10	HSRPTK15
25	241.4	6.6	768	12	AF016247
26	236.2	6.5	1229	10	HSRPTK123
27	226.4	6.2	661	10	HSRPTK14
28	218	6.0	400	14	G11190
29	200	5.5	2738	4	GGU43396
30	198.4	5.5	1568	10	HSTRK2H
31	196.8	5.4	2727	4	GGTRKA
32	196.8	5.4	2010	10	HSTRK13ON
33	193.6	5.3	2301	5	I96186
34	193.6	5.3	2701	9	HUMTRKPOA
35	193.6	5.3	2701	9	HSTRKR
36	193.6	5.3	2301	10	HSTRK1
37	193.6	5.3	1746	10	HSTRK1
38	190	5.2	3707	5	I35538
39	190	5.2	3194	10	S76473
40	185.8	5.1	2697	12	HSU12140
41	185.8	5.1	2682	12	RATTRKCA
42	185.8	5.1	2682	12	RATTRKCN3
43	185.8	5.1	2697	12	S62924
44	185.2	5.1	1408	11	AF052184
45	184.4	5.1	2067	36	GCTK

ALIGNMENTS

RESULT	1
I80845	
LOCUS	I80845
DEFINITION	Sequence 3 from patent US 5709858.
ACCESSION	I80845
NID	93209135
VERSION	I80845.1
KEYWORDS	GI:3209135
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	1 (bases 1 to 3637)
AUTHORS	Godowski, P.J., Mark, M.R. and Scadden, D.T.
TITLE	Antibodies specific for Rse receptor protein tyrosine kinase
JOURNAL	Patent: US 5709858-A 3 20-JAN-1998;
FEATURES	Location/Qualifiers
source	1..3637
BASE COUNT	721 a 1094 c 1079 g 743 t
ORIGIN	

Query Match 100.0%; Score 3637; DB 5; Length 3637;  
Best Local Similarity 100.0%; Pred. No. 0;

Matches 3637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCGAGTCGACGTTGGACCTTGAAGGAATGCAAGAGATGCTGCCCCACACCCCTT 60  
Db 1 GAATTCGAGTCGACGTTGGACCTTGAAGGAATGCAAGAGATGCTGCCCCACACCCCTT 60

QY 61 AGGCCCCGAGGATCAGGAGCTATGGACACAGAGCCCTGTCACTTTACTGCTGCTC 120  
Db 61 AGGCCCCGAGGATCAGGAGCTATGGACACAGAGCCCTGTCACTTTACTGCTGCTC 120

QY 121 TTGGTGGCAAGTGGAGATGCTCACATGAAGGACATTTTGATCTGCTGCAAGTCCGCTAT 180  
Db 121 TTGGTGGCAAGTGGAGATGCTCACATGAAGGACATTTTGATCTGCTGCAAGTCCGCTAT 180

QY 181 GCCCTGGGCATCAGAGCCGACCATCCAGACAGTACATCTCTGCTTCCAGTCTCTGG 240  
Db 181 GCCCTGGGCATCAGAGCCGACCATCCAGACAGTACATCTCTGCTTCCAGTCTCTGG 240

QY 241 TCAGATTCACCTGCGCCCGCCACAGAGTTGGAGACAGTACGAGGAGTGGGCGCTGG 300  
Db 241 TCAGATTCACCTGCGCCCGCCACAGAGTTGGAGACAGTACGAGGAGTGGGCGCTGG 300

QY 301 TGCCCGCAGGCTCGGTGTTCCCAAGCAGGAGGAGTACTTGCAGGTGGATCTACAACGA 360  
Db 301 TGCCCGCAGGCTCGGTGTTCCCAAGCAGGAGGAGTACTTGCAGGTGGATCTACAACGA 360

QY 361 CTGACCTGCTGGTCTGCTGGGACACGAGGCGCATGCCGGGCGCTGGGCAAGGAG 420  
Db 361 CTGACCTGCTGGTCTGCTGGGACACGAGGCGCATGCCGGGCGCTGGGCAAGGAG 420

QY 421 TTCTCCCGAGCTACCGGCTCGGTACTTCCCGGATGGTCCCGCTGGATGGGCTGAAG 480  
Db 421 TTCTCCCGAGCTACCGGCTCGGTACTTCCCGGATGGTCCCGCTGGATGGGCTGAAG 480

QY 481 GACCGCTGGGTGAGGAGTATCTAGGCAATGAGGACCTGAGGAGTGGTCTCAAG 540  
Db 481 GACCGCTGGGTGAGGAGTATCTAGGCAATGAGGACCTGAGGAGTGGTCTCAAG 540

QY 541 GACCTGGGCCCCCGATGGTTCGCGACTGTTCCGCTTACCCCCGGGTGACCGGGTC 600  
Db 541 GACCTGGGCCCCCGATGGTTCGCGACTGTTCCGCTTACCCCCGGGTGACCGGGTC 600

QY 601 ATGAGCGCTCTGTCGGGTGAGCTCTATGGCTGCTCTGGAGGATGACTCTGCT 660  
Db 601 ATGAGCGCTCTGTCGGGTGAGCTCTATGGCTGCTCTGGAGGATGACTCTGCT 660

QY 661 TACACCGCCCCCTGTGGGCGACACAATGATTTATCTCAGGCGCTGTACCTCAAGACTCC 720  
Db 661 TACACCGCCCCCTGTGGGCGACACAATGATTTATCTCAGGCGCTGTACCTCAAGACTCC 720

QY 721 ACCTATGACGACATACCGTGGGCGGACTCGAGTATGGGGTCTGGCCAGCTGGCAGAT 780  
Db 721 ACCTATGACGACATACCGTGGGCGGACTCGAGTATGGGGTCTGGCCAGCTGGCAGAT 780

QY 781 GGTGTGTGGGCTGGATGACTTTAGGAAGTACAGAGCTGGGGTCTGGCCAGCTAT 840  
Db 781 GGTGTGTGGGCTGGATGACTTTAGGAAGTACAGAGCTGGGGTCTGGCCAGCTAT 840

QY 841 GACTATGTGGATGGACCAACACAGCTTCTCCAGTGGCTATGTGGAGATGGAATTTAG 900  
Db 841 GACTATGTGGATGGACCAACACAGCTTCTCCAGTGGCTATGTGGAGATGGAATTTAG 900

QY 901 TTTCACCGGCTGAGGGCTTCCAGGCTATCGAGTCCAGTGTAAACATGACACGCTG 960  
Db 901 TTTCACCGGCTGAGGGCTTCCAGGCTATCGAGTCCAGTGTAAACATGACACGCTG 960

QY 961 GGAGCCGCTGCTGCTGGGGGTGGAATGTCGCTTCCGGGCTGGCCCTGCGATGGCTGG 1020  
Db 961 GGAGCCGCTGCTGCTGGGGGTGGAATGTCGCTTCCGGGCTGGCCCTGCGATGGCTGG 1020

QY 1021 GAGGGGAGCCCATGCGCCACACCTAGGGGGCAACCTGGGGGACCCCGAGCCGGCT 1080  
Db 1021 GAGGGGAGCCCATGCGCCACACCTAGGGGGCAACCTGGGGGACCCCGAGCCGGCT 1080

QY 1081 GTCTAGTGCCTTGGCGCGCTGGCTCGCTTTCTGAGTGGCGCTTCTCTTTGG 1140  
Db 1081 GTCTAGTGCCTTGGCGCGCTGGCTCGCTTTCTGAGTGGCGCTTCTCTTTGG 1140

QY 1141 GGCCCTGTGTTACTCTTCAGCGAAATCTCTCTATCTGATGTGTGAACAAATCTCTCT 1200  
Db 1141 GGCCCTGTGTTACTCTTCAGCGAAATCTCTCTATCTGATGTGTGAACAAATCTCTCT 1200

QY 1201 CCGGCACTGGAGGCACTTCCCGCCAGCCCTGTGTGGCCGCTGGCCAGCCCTCCAC 1260  
Db 1201 CCGGCACTGGAGGCACTTCCCGCCAGCCCTGTGTGGCCGCTGGCCAGCCCTCCAC 1260

QY 1261 AACTTCAGCAGCTTGGAGCTGGAGCCAGAGCCAGCAGCCGTCGAGCCGTCGAGGGG 1320  
Db 1261 AACTTCAGCAGCTTGGAGCTGGAGCCAGAGCCAGCAGCCGTCGAGCCGTCGAGGGG 1320

QY 1321 AGCCGACCGCATCTCTATCGGCTGCTGTGGCATCATCTGCTGCTGCTGCTCATC 1380  
Db 1321 AGCCGACCGCATCTCTATCGGCTGCTGTGGCATCATCTGCTGCTGCTGCTCATC 1380

QY 1381 ATTGCCCTCATGCTGTGGCGGTGACTGGCCAGGCTCTCAGCAAGGCTGAACGGAG 1440  
Db 1381 ATTGCCCTCATGCTGTGGCGGTGACTGGCCAGGCTCTCAGCAAGGCTGAACGGAG 1440

QY 1441 GTGTTGGAAGAGAGCTGACGTTTCACTCTCTGCTGCTGGGACACTATCTCTATCAAC 1500  
Db 1441 GTGTTGGAAGAGAGCTGACGTTTCACTCTCTGCTGCTGGGACACTATCTCTATCAAC 1500

QY 1501 AACCGCCAGGTCTTAGAGAGCCACCCCGTACAGAGCCCGGCGCTCGTGGGAATCCG 1560  
Db 1501 AACCGCCAGGTCTTAGAGAGCCACCCCGTACAGAGCCCGGCGCTCGTGGGAATCCG 1560

QY 1561 CCGCACTCCGCTCCCTGTGCTCCCAATGGCTGCTGCTGCTCTCCAATCCAGCTAC 1620  
Db 1561 CCGCACTCCGCTCCCTGTGCTCCCAATGGCTGCTGCTGCTCTCCAATCCAGCTAC 1620

QY 1621 CGCCCTCTTCTGGCCACTTACCGCGCTCCCGCTCGAGCGCCGCGGCCCCCAGCCGCC 1680  
Db 1621 CGCCCTCTTCTGGCCACTTACCGCGCTCCCGCTCGAGCGCCGCGGCCCCCAGCCGCC 1680

QY 1681 TGGGCAAAACCCACCAACAGCCAGGCTACAGTGGGACTATATGAGGCTTGAGAAGCA 1740  
Db 1681 TGGGCAAAACCCACCAACAGCCAGGCTACAGTGGGACTATATGAGGCTTGAGAAGCA 1740

QY 1741 GCGCCCCGCTTCTGCCCCACCTCCCGAGACAGCTGCCCCCATTTATGCGGAGGCTGAC 1800  
Db 1741 GCGCCCCGCTTCTGCCCCACCTCCCGAGACAGCTGCCCCCATTTATGCGGAGGCTGAC 1800

QY 1801 ATTTGTTACCCCTGAGGCGCTCACCGGGGCAACACCTATGCTGCTGCTGCTGCTGCTG 1860  
Db 1801 ATTTGTTACCCCTGAGGCGCTCACCGGGGCAACACCTATGCTGCTGCTGCTGCTGCTG 1860

QY 1861 GGGGAGCTGGGGATGGGCCCCCGAGAGTGAATTTCCCTCGACTCGGCTTCAAG 1920  
Db 1861 GGGGAGCTGGGGATGGGCCCCCGAGAGTGAATTTCCCTCGACTCGGCTTCAAG 1920

QY 1921 GAGAAGCTTGGCAGGCGCTGTTGGGAGGTGACCTGCTGAGGTGCGAGAGCCCTCAA 1980  
Db 1921 GAGAAGCTTGGCAGGCGCTGTTGGGAGGTGACCTGCTGAGGTGCGAGAGCCCTCAA 1980

QY 1981 GATCTGTGCTGTTGATTTCCCTTAAATGTGGGTAAAGGACACCTTTGCTGGTAGCT 2040  
Db 1981 GATCTGTGCTGTTGATTTCCCTTAAATGTGGGTAAAGGACACCTTTGCTGGTAGCT 2040

QY 2041 GTCAGATCTTACGCCAGATGCGCCAGAAATGCCAGAAATGATTTCTGAAAGAGGTG 2100  
Db 2041 GTCAGATCTTACGCCAGATGCGCCAGAAATGCCAGAAATGATTTCTGAAAGAGGTG 2100

QY 2101 AAGATCATGTGAGGCTCAAGGACCCCAACATCATTCGGCTGCTGGGCTGTGTGTCAG 2160  
Db 2101 AAGATCATGTGAGGCTCAAGGACCCCAACATCATTCGGCTGCTGGGCTGTGTGTCAG 2160









```
Db 2357 TCAGCTACCAATGCTGCTGATGTGGACGCCACGATGCCCTCCGGCATGCGCTATCTGG 2416
QY 2354 CCACACTCAACTTTGACATCGGACCTGGCCACCGGAACCTGCTAGTTGGGAAATTT 2413
Db 2417 CCACACTCAACTTTGACATCGGACCTGGCCACCGGAACCTGCTAGTTGGGAAATTT 2416
QY 2414 TCACCATCAAAATCGCAGACTTTGGCATGAGCGGAACTCTATGCTGGGACATATACC 2473
Db 2477 TCACCATCAAAATCGCAGACTTTGGCATGAGCGGAACTCTATGCTGGGACATATACC 2536
QY 2474 GTGTGAGGCGCGGACGCTGCTGCCATCCGCTGGATGGCTGGAGTGCATCCTCATGG 2533
Db 2537 GTGTGAGGCGCGGACGCTGCTGCCATCCGCTGGATGGCTGGAGTGCATCCTCATGG 2536
QY 2534 GGAAGTTACAGACTGCGAGTGCAGTGTGGGCGCTTTGGTGTGACCCCTGTGGAGGTGCTGA 2593
Db 2597 GGAAGTTACAGACTGCGAGTGCAGTGTGGGCGCTTTGGTGTGACCCCTGTGGAGGTGCTGA 2656
QY 2594 TGCCTGTAGGCGCCAGCCCTTTGGCAGCTCACCGAGGACGAGTGCATCAGAACCGGG 2653
Db 2657 TGCCTGTAGGCGCCAGCCCTTTGGCAGCTCACCGAGGACGAGTGCATCAGAACCGGG 2716
QY 2654 GGGAGTTCTTCGGGACCGAGCGCGGAGGTGACTGTCCGGCGCGCTGCCTGCCCGC 2713
Db 2717 GGGAGTTCTTCGGGACCGAGCGCGGAGGTGACTGTCCGGCGCGCTGCCTGCCCGC 2776
QY 2714 AGGGCTATATGAGTGTGCTTCGGTGTGGGCGGAGTCTGAGCAGCGACCAACCT 2773
Db 2777 AGGGCTATATGAGTGTGCTTCGGTGTGGGCGGAGTCTGAGCAGCGACCAACCT 2836
QY 2774 TTTCCAGCTGATGGTCTTCGGCAGGAGTGCATCAACACGCTGTGATCAGACATC 2833
Db 2837 TTTCCAGCTGATGGTCTTCGGCAGGAGTGCATCAACACGCTGTGATCAGACATC 2896
QY 2834 CAGCTGCCCTCCCTCAGGAGTGTACAGGGAAGCAGTGCATCAACACAGAGGAC 2893
Db 2897 CAGCTGCCCTCCCTCAGGAGTGTACAGGGAAGCAGTGCATCAACACAGAGGAC 2956
QY 2894 ACAATGGACCTCTG-CGCTTCCCTCCGACAGCCCATCCTCTAATAGAGCAGTGA 2952
Db 2957 ACAATGGACCTCTGCGCCCTTCCCTCCGACAGCCCATCCTCTAATAGAGCAGTGA 3016
QY 2953 GACTGC----- 2959
Db 3017 GACTGCAGGCTGGGCGCCACCCAGGAGTGTAGCCCTTCTCCCTTCTCTGGACACATC 3076
QY 2959 -----AGAGCCCTCTGTGCCACCCAGCTGTCTGTGCCACCCAGCTGTCTGT 2990
Db 3077 TCATGTCCCTTCTGTCTTCTTCTTAGAGCCCTGTGTGCCACCCAGCTGTCTGT 3136
QY 2991 TGGATGGATCTCTCCAGCTCTCTAGCCATCCCTTTGGGGAAGGTGGGGAATAAT 3050
Db 3137 TGGATGGATCTCTCCAGCTCTAGCCATCCCTTTGGGGAAGGTGGGGAATAAT 3196
QY 3051 AGGATAGACACTGGACATGGCCCATTTGAGACCTTGGGCCCCCAGCTGATTC 3110
Db 3197 AGGATAGACACTGGACATGGCCCATTTGAGACCTTGGGCCCCCAGCTGATTC 3256
QY 3111 CTGGAGAGTGTGCTGGGCGCCAGCTTCTCTCTCCCTGTTCACACTGGACCCCACTGGC 3170
Db 3257 CTGGAGAGTGTGCTGGGCGCCAGCTTCTCTCTCCCTGTTCACACTGGACCCCACTGGC 3316
QY 3171 TGAGAACTGGGGGTGAGGAGACAGAGGAGGAAATTTTCTTGTGCTGCTGCC 3230
Db 3317 TGAGAACTGGGGGTGAGGAGACAGAGGAGGAAATTTTCTTGTGCTGCTGCC 3376
QY 3231 TGTACTTGTCTCAGTTGGGCTTCTTCTCTCTCCATCCTGAAACACTGGACCTGGGG 3290
Db 3377 TGTACTTGTCTCAGTTGGGCTTCTTCTCTCTCCATCCTGAAACACTGGACCTGGGG 3436
QY 3291 GTAGCCCGCCCGCCAGCTCTAGTCA-CCGCCACTTCCACCTGCGAGTCTTCTAGTAGAA 3349
Db 3437 GTAGCCCGCCCGCCAGCTCTAGTCAACCCCACTTCCACCTGCGAGTCTTCTAGTAGAA 3496
```

```
QY 3350 CTTCTTAAGCCTATAGCTTTCTGTGAGTAATAATTGGATTGGGGGAAAGAGGAGC 3409
Db 3497 CTTCTTAAGCCTATAGCTTTCTGTGAGTAATAATTGGATTGGGGGAAAGAGGAGC 3556
QY 3410 AAGCGCCCATAGCCTTGGGTTGGACATCTCTAGTCTAGCTGCCACATGATTTTCTAT 3469
Db 3557 AAGCGCCCATAGCCTTGGGTTGGACATCTCTAGTCTAGCTGCCACATGATTTTCTAT 3616
QY 3470 AATCACTTGGGTTTGTACATTTTGGGGGAGACACAGATTTTACACATAATATG 3529
Db 3617 AATCACTT-GGGTTTGTACATTTTGGGGGAGACACAGATTTTACACATAATATG 3675
QY 3530 GACCTAGCTTGAAGCAATTTTAAATCCCTGCCTAGCAGGTAATAATAAGGTTGAGTT 3589
Db 3676 GACCTAGCTTGAAGCAATTTTAAATCCCTGCCTAGCAGGTAATAATAAGGTTGAGTT 3735
QY 3590 TTCACACAAAAA 3608
Db 3736 TTCACACAAAAA 3754

RESULT 4
HUMRTK HUMRTK 3738 bp mRNA PRI 27-SEP-1993
LOCUS Homo sapiens receptor tyrosine kinase mRNA, complete cds.
DEFINITION L11315
ACCESSION 9403386
NID L11315.1 GI:403386
VERSION receptor tyrosine kinase.
KEYWORDS Homo sapiens (library: lambda-gtl1) term placenta cDNA to mRNA.
SOURCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
ORGANISM Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 3738)
AUTHORS Johnson,J.D., Edman,J.C. and Rutter,W.J.
TITLE A receptor tyrosine kinase found in breast carcinoma cells has an
extracellular discoidin I-like domain
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 90, 5677-5681 (1993)
MEDLINE 93296201
COMMENT On Sep 27, 1993 this sequence version replaced gi:307397.
FEATURES
Location/Qualifiers
source
1..3738
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="term placenta"
/tissue_lib="lambda-gtl1"
1..141
142..2883
/standard_name="discoidin I-like domain receptor tyrosine
kinase"
/codon_start=1
/product="receptor tyrosine kinase"
/protein_id="AA02866.1"
/db_xref="PIR:G403387"
/db_xref="GI:403387"
/translation="MGPEALSSLLLLVSGADMDKGFDPKAKRYALGMODRTIPD
SDISASSSDSTAARHSRLESSDGDGACPGSVFPEEEYLOVDLQRLHVALVGT
QQRHAGGLKEFSRYRLRYSRDGRWGWDRGQEVISNEDPEGVLKDLGPPMV
ARLRFYPRADRVMSVLRVELYGLCLRWLLSYTAPVQTMYLEAYLINDYDGH
TVGGLOYGLGQLADGVVGLDFRKSQELRWPGDYVGVNHSFSSGYVEMEFDR
LRAQAMOVHCNNMHTLGAIRPGVECFRRGPMAMWEGEPMRHNLGNLGPRAVR
SVPLGRVARFLQCRFLFAGPWLFLSEISFTSDVNNSSPALGGTFPPAPWPGPPP
TNFSSLEPEPGQOPVAKAGSPAILIGCLVAITLLLLLIALLMLRHLRLLSKA
ERXLEELTVHLSVPGDTILINRRPREPPYOEPNRPNGPSPFCVNGSALLL
SNAPRLLLATYAPPPGPTPAWAKPTNTQYSGDMEPEKPGAPLLPPQNSV
PHYAADIIVTQGVGTNTYAVPALPGAVGDPVPDFPRSLRFLKELGQFGEV
HLCFVSDQLVSLDFPLNVRKHPHLLVAVLKLPRDANKRNDFLKVKTMSRLKP
NTIRLLGVCVDDPLCLMTDYMGDLNQFLSAHOLEDKAAEGAPCDGQAAGTISY
PMLHVAQIASGMEYLATNFRHDLATNCLNGENTIKIADFGRSNRIYAGIYR
VOGRAVLPIRMHAWECILMGKFTTASDVWAFGVTVWVEMILCRAQPFQGLFDEQVIE
AGEFFRDQGRQVYLSRRPACPGQLYELMLRCWSESEQRPFPSQLHRELAEDALNTV"
2887..3738
3'UTR
```







Db 2117 CAGATGCCAACAAATGCGAGGAATCATTTCTGAAAGAGGTGAAGATCATGCTCGAGGC 2176  
Qy 2117 TCAGAGGCCAAACATCATTTGCGGTGCTGGCGGTGTGTGTCAGAGGAGGACCCCTCTGCA 2176  
Db 2177 TCAAGGACCCAAACATCATTTGCGGTGCTGGCGGTGTGTGTCAGAGGAGGACCCCTCTGCA 2236  
Qy 2177 TGATTACTGACTACATGGAGACGGCCACTCAACCACTGTTCTCAGTGGCCACCACTGG 2236  
Db 2237 TGATTACTGACTACATGGAGACGGCCACTCAACCACTGTTCTCAGTGGCCACCACTGG 2296  
Qy 2237 AGGACAAAGGACGGCCAGGCGCCCTGGGACGGCCAGGCTGCGGACGGCCATCA 2296  
Db 2297 AGGACAAAGGACGGCCAGGCGCCCTGGGACGGCCAGGCTGCGGACGGCCATCA 2356  
Qy 2297 GCTACCCAAATGCTGCTGATGTGCGAGCCAGATCGCTCGGCGATGCTATCTGCGCA 2356  
Db 2357 GCTACCCAAATGCTGCTGATGTGCGAGCCAGATCGCTCGGCGATGCTATCTGCGCA 2416  
Qy 2357 CACTCAACTTTGTATCGGGACCTGGCCACGGCAACTGCTAGTTGGGAAATTTCA 2416  
Db 2417 CACTCAACTTTGTATCGGGACCTGGCCACGGCAACTGCTAGTTGGGAAATTTCA 2476  
Qy 2417 CCATCAAAATCCGAGACTTTGGCATGAGCCGGAACCTCTATGCTGGGACTATTACCGTG 2476  
Db 2477 CCATCAAAATCCGAGACTTTGGCATGAGCCGGAACCTCTATGCTGGGACTATTACCGTG 2536  
Qy 2477 TGCAGGCGCGGAGTGTGCTGCCATCCGCTGGATGCGCTGGGAGTGCATCTCATGGGA 2536  
Db 2537 TGCAGGCGCGGAGTGTGCTGCCATCCGCTGGATGCGCTGGGAGTGCATCTCATGGGA 2596  
Qy 2537 AGTTCACCACTGCGAGTGTGCGGCTTTGGTGTGACCGCTGGGAGTGTGATGC 2596  
Db 2597 AGTTCACCACTGCGAGTGTGCGGCTTTGGTGTGACCGCTGGGAGTGTGATGC 2656  
Qy 2597 TCTGTAGGCGCCAGCCCTTTGGCAGCTTCACCGAGCAGGTCTGAGAACCGCGGGG 2656  
Db 2657 TCTGTAGGCGCCAGCCCTTTGGCAGCTTCACCGAGCAGGTCTGAGAACCGCGGGG 2716  
Qy 2657 AGTTCCTCCGGACCGAGGCGGAGGTGTACCTGTCCCGCGGCTGCTGCGCGGAGG 2716  
Db 2717 AGTTCCTCCGGACCGAGGCGGAGGTGTACCTGTCCCGCGGCTGCTGCGCGGAGG 2776  
Qy 2717 GCCTATATGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2776  
Db 2777 GCCTATATGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2836  
Qy 2777 CCAGCTGCATCGGTTCTTGGCAGAGGATGCTCAACACCGGTGTGATCATACATCCAG 2836  
Db 2837 CCAGCTGCATCGGTTCTTGGCAGAGGATGCTCAACACCGGTGTGATCATACATCCAG 2896  
Qy 2837 CTGCGCCCTCCCTCAGGAGTGTATCAGGGGAGCCAGTGCACCTAAACACAGAGGACACA 2896  
Db 2897 CTGCGCCCTCCCTCAGGAGGAGTGTATCAGGGGAGCCAGTGCACCTAAACACAGAGGACACA 2956  
Qy 2897 ATGGCAGCTCTG-CCCTTCCCTCCCGACAGCCCATCACTCACTTAATAGGACGAGTGCAG 2955  
Db 2957 ATGGCAGCTCTG-CCCTTCCCTCCCGACAGCCCATCACTCACTTAATAGGACGAGTGCAG 3016  
Qy 2956 TGC----- 2959  
Db 3017 TGCAGGCTGGGCCACCCAGGAGGTGATGCCCTTCTCCCTTCTGACACACTCTCA 3076  
Qy 2959 -----AGAGCCCTGTGCCCCACCCAGCTGCTGTGG 2993  
Db 3077 TGTCCCTTCTCTGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3136  
Qy 2994 ATGGGATCTCTCCACCTCTCTAGCCATCCCTTGGGAGGCTGGGAGGAAATATAGG 3053  
Db 3137 ATGGGATCTCTCTCCACCTCTCTAGCCATCCCTTGGGAGGCTGGGAGGAAATATAGG 3196  
Qy 3054 ATAGACACTGGACATGGCCCATTTGGACACCTTGGGCCCCACTGGACAACTGATTCCTG 3113

Db 3197 ATAGACACTGGACATGGCCCATTTGGAGCACCTGGGCCCACTGGACAACTGATTCCTG 3256  
Qy 3114 GAGAGGTGGCTGGCCCGCCAGCTTCTCTCTCCCTGTACACACTGGACCCCACTGGGTGA 3173  
Db 3257 GACAGGTGGCTGGCCCGCCAGCTTCTCTCTCCCTGTACACACTGGACCCCACTGGGTGA 3316  
Qy 3174 GAATCTGGGGTGGAGGAGCAAGAAGAGGAGGAAATGTTTCTTGTGCTGCTGCTGCT 3233  
Db 3317 GAATCTGGGGTGGAGGAGCAAGAAGAGGAGGAAATGTTTCTTGTGCTGCTGCTGCT 3376  
Qy 3234 ACTTGTCTCAGCTTGGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3293  
Db 3377 ACTTGTCTCAGCTTGGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3436  
Qy 3294 GCGCGCGCCCGCCCTCAGTCA-CCCCCACTTCCCACTGCGAGTCTTCTAGTAGTAACTT 3352  
Db 3437 GCGCGCGCCCGCCCTCAGTCA-CCCCCACTTCCCACTGCGAGTCTTCTAGTAGTAACTT 3496  
Qy 3353 CTCTAAGCCTATACGTTTCTTGTGGAGTAAATATTGGGATTTGGGGGAAAGAGGAGCAAC 3412  
Db 3497 CTCTAAGCCTATACGTTTCTTGTGGAGTAAATATTGGGATTTGGGGGAAAGAGGAGCAAC 3556  
Qy 3413 GCGCATAGCCTTGGGTTGGACATCTCTAGTGTAGCTGCCACATTTTCTTATTAAT 3472  
Db 3557 GCGCATAGCCTTGGGTTGGACATCTCTAGTGTAGCTGCCACATTTTCTTATTAAT 3616  
Qy 3473 CACTTGGGTTTGTACATTTTGGGGGAGAGACACAGATTTTACATAATATGAC 3532  
Db 3617 CACTT-GGGTTTGTACATTTTGGGGGAGAGACACAGATTTTACATAATATGAC 3675  
Qy 3533 CTAGCTTGGAGCAATTTAATCCCTGCTAGGAGGATTAATAAGGTTGAGTTTTC 3592  
Db 3676 CTAGCTTGGAGCAATTTAATCCCTGCTAGGAGGATTAATAAGGTTGAGTTTTC 3735  
Qy 3593 CAC 3595  
Db 3736 CAC 3738

## RESULT 5

HUMCAK HUMCAK 3803 bp mRNA PRI 11-MAY-1994  
LOCUS Homo sapiens tyrosine protein kinase (CAK) gene, complete cds.  
DEFINITION L20817  
ACCESSION L20817  
VERSION 9306474  
KEYWORDS L20817.1 GI:306474  
SOURCE A5-antigen; factor V; factor VIII; tyrosine protein kinase.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 3803)  
AUTHORS Perez, J.L., Shen, X., Finkernagel, S., Sciorra, L., Jenkins, N.A.,  
Gilbert, D.J., Copeland, N.G. and Wong, T.W.  
TITLE Identification and chromosomal mapping of a receptor tyrosine  
kinase with a putative phospholipid binding sequence in its  
ectodomain  
JOURNAL Oncogene 9, 211-219 (1994)  
MEDLINE 9413417  
REFERENCE 2 (bases 1 to 3803)  
AUTHORS Johnson, J.D.  
TITLE Direct Submission  
JOURNAL Submitted (17-FEB-1993) J.D. Johnson  
REFERENCE 3 (sites)  
AUTHORS Johnson, J.D., Edman, J.C. and Rutter, W.J.  
TITLE A receptor tyrosine kinase found in breast carcinoma cells has an  
extracellular discoidin I-like domain  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 90, 5677-5681 (1993)  
MEDLINE 93296201  
AUTHORS Wong, T.  
TITLE Direct Submission  
JOURNAL Submitted (06-JUL-1993) T. Wong



Db 1589 CTGGGACACTATCTCATCAACAACCGCCAGGTCCTTAGAGAGCCACCCCGGTACAGG 1648  
Qy 1538 AGCCCGGCTCGTGGGAATCCGCCACTCCGCTCCCTGTGTGCCCAATGGCTCTGGGT 1597  
Db 1649 AGCCCGGCTCGTGGGAATCCGCCACTCCGCTCCCTGTGTGCCCAATGGCTCTGGGT 1708  
Qy 1598 TGTGCTCTCGAATCAGGCTACGGCTTCCTTCTGGCCACTTTACGCCGTCCTCCCTCGAG 1657  
Db 1709 TGTGCTCTCGAATCAGGCTACGGCTTCCTTCTGGCCACTTTACGCCGTCCTCCCTCGAG 1768  
Qy 1658 GCCCGGGCCCCCACACCCGCTGGGCCAAACCAACACCCAGCCGCTACAGTGGGG 1717  
Db 1769 GCCCGGGCCCCCACACCCGCTGGGCCAAACCAACACCCAGCCGCTACAGTGGGG 1828  
Qy 1718 ACTATATGAGGCTGAGAAAGCCAGCGCCCGCTTCTGCCCCACCTCCCCAGAACAGCG 1777  
Db 1829 ACTATATGAGGCTGAGAAAGCCAGCGCCCGCTTCTGCCCCACCTCCCCAGAACAGCG 1888  
Qy 1778 TCCCCCATTTATGCGAGGCTGACATTTTACCTCTGAGGGGCTCACCGGGGGCAACACCT 1837  
Db 1889 TCCCCCATTTATGCGAGGCTGACATTTTACCTCTGAGGGGCTCACCGGGGGCAACACCT 1948  
Qy 1838 ATGCTGTGCTGCTGACCTGCCCCAGGGGAGTCGGGATGGGCCCCCCCAGAGTGGATTCC 1897  
Db 1949 ATGCTGTGCTGCTGACCTGCCCCAGGGGAGTCGGGATGGGCCCCCCCAGAGTGGATTCC 2008  
Qy 1898 CTGATCTCGACTCCGCTTCAAGGAGAACTTTGGGAGGGCCAGTTTGGGAGGTGCACC 1957  
Db 2009 CTGATCTCGACTCCGCTTCAAGGAGAACTTTGGGAGGGCCAGTTTGGGAGGTGCACC 2068  
Qy 1958 TGTGTGAGTGCAGACCCCTCAAGATCTGGTCAGTCTTGATTTCCCTTAAATGTGCGTA 2017  
Db 2069 TGTGTGAGTGCAGACCCCTCAAGATCTGGTTAGTCTTGATTTCCCTTAAATGTGCGTA 2128  
Qy 2018 AGGAGACCCCTTGTCTGTGTGCTGCTCAAGATCTTACGGCCAGATGCCACCAAGATGCCA 2077  
Db 2129 AGGAGACCCCTTGTCTGTGTGCTGCTCAAGATCTTACGGCCAGATGCCACCAAGATGCCA 2188  
Qy 2078 GGAATGATTTCTTGAAGAGGTGAAGATCATGTGAGGCTCAAGGACCCCAACATCATTC 2137  
Db 2189 GGAATGATTTCTTGAAGAGGTGAAGATCATGTGAGGCTCAAGGACCCCAACATCATTC 2248  
Qy 2138 GGTGCTGGGCTGTGTGTGAGGAGGACCCCTCTGCAATGATTTACTGACTACATGGAGA 2197  
Db 2249 GGTGCTGGGCTGTGTGTGAGGAGGACCCCTCTGCAATGATTTACTGACTACATGGAGA 2308  
Qy 2198 ACGGCAGCTCAACCACTTCCCTCAGTGCCACACAGCTGGAGGACAGGACCGGAGGGG 2257  
Db 2309 ACGGCAGCTCAACCACTTCCCTCAGTGCCACACAGCTGGAGGACAGGACCGGAGGGG 2368  
Qy 2258 CCCTGGGACGGCAGGCTGGGAGGGGCCACCATCAGCTACCCATGCTGTGCTGCTG 2317  
Db 2369 CCCTGGGACGGCAGGCTGGGAGGGGCCACCATCAGCTACCCATGCTGTGCTGCTG 2428  
Qy 2318 TGGAGCCAGATCGGCTCCGGCATCGGTATCTGCGCCACATCACTTTGTACATCGGG 2377  
Db 2429 TGGAGCCAGATCGGCTCCGGCATCGGTATCTGCGCCACATCACTTTGTACATCGGG 2488  
Qy 2378 ACCTGGCCCGGCACTGCCTAGTTTGGGAAATTTACCATCAAAATCGCAGACTTTG 2437  
Db 2489 ACCTGGCCCGGCACTGCCTAGTTTGGGAAATTTACCATCAAAATCGCAGACTTTG 2548  
Qy 2438 GCATAGCGGAACCTCTATGCTGGGACTATTACCGTGTGGAGGCGGGCAGTGTGCTG 2497  
Db 2549 GCATAGCGGAACCTCTATGCTGGGACTATTACCGTGTGGAGGCGGGCAGTGTGCTG 2608  
Qy 2498 CCATCCGCTGGATGGCTGGAGTGCATCCTCATGGGAAATTTACGACTGCGAGTGGAG 2557  
Db 2609 CCATCCGCTGGATGGCTGGAGTGCATCCTCATGGGAAATTTACGACTGCGAGTGGAG 2668  
Qy 2558 TGTGGGCTTTGGTGTGACCCCTGTGGAGGTGCTGATGCTGTAGGGCCGACCCCTTTG 2617  
Db 2669 TGTGGGCTTTGGTGTGACCCCTGTGGAGGTGCTGATGCTGTAGGGCCGACCCCTTTG 2728

Qy 2618 GGAGGTCAACGACGAGGTCATCGAAGAACGCGGGGAGTTCTTCCGGGACCCAGGGCC 2677  
Db 2729 GGAGGTCAACGACGAGGTCATCGAAGAACGCGGGGAGTTCTTCCGGGACCCAGGGCC 2788  
Qy 2678 GGAGGTGTACCTGTCCCGCCGCTGCTGCGCCGAGGCGCTATATGAGCTGATGCTTC 2737  
Db 2789 GGAGGTGTACCTGTCCCGCCGCTGCTGCGCCGAGGCGCTATATGAGCTGATGCTTC 2848  
Qy 2738 GGTGCTGAGCCCGGAGTCTGAGCAGCACACCTTTTCCAGCTGCACTGGTTCCTGG 2797  
Db 2849 GGTGCTGAGCCCGGAGTCTGAGCAGCACACCTTTTCCAGCTGCACTGGTTCCTGG 2908  
Qy 2798 CAGAGATGACACTCAACACGCTGTGAATCACATCAGCTGCCCTCTCCCTCAGGAGTG 2857  
Db 2909 CAGAGATGACACTCAACACGCTGTGAATCACATCAGCTGCCCTCTCCCTCAGGAGTG 2968  
Qy 2858 ATCCAGGGAAGCCAGTGCACACTAAACAGAGAGACACATGCGACCTCTGCCCTTCCCC 2917  
Db 2969 ATCCAGGGAAGCCAGTGCACACTAAACAGAGAGACACATGCGACCTCTGCCCTTCCCC 3028  
Qy 2918 TCCGACAGCCCATCAGCTCTAATAGAGGAGTGAGACTGC ----- 2959  
Db 3029 TCCGACAGCCCATCAGCTCTAATAGAGGAGTGAGACTGCAGGTGGCTGGGCAAGGA 3088  
Qy 2959 ----- 2959  
Db 3089 GCTGATGCCCTTCTCCCTTCTGGACACACTCTCATGTGCCCTTCTTCTTCTTCTTC 3148  
Qy 2959 --AGAAGCCCTGTGCGCCACAGCTGCTGTGTGATGGGATCTTCCACCCCTCTC 3016  
Db 3149 CTAGAAGCCCTGTGCGCCACAGCTGCTGTGTGATGGGATCTTCCACCCCTCTC 3208  
Qy 3017 TAGCCATCCCTTGGGGAAGGTGGGAGAAATATAGATAGACACTGGACATGCCCCATT 3076  
Db 3209 TAGCCATCCCTTGGGGAAGGTGGGAGAAATATAGATAGACACTGGACATGCCCCATT 3268  
Qy 3077 GGAGCAGCTGGGCCCCACTGGCAACACTGATTCTCTGGAGAGTGGCTGCCGCCAGCT 3136  
Db 3269 GGAGCAGCTGGGCCCCACTGGCAACACTGATTCTCTGGAGAG -----GAGACCCAGCT 3322  
Qy 3137 TCTCTCTCCCTGTACACACTGGACCCCACTGGCTGAGAAATCTGGGGGTGAGGAGACAA 3196  
Db 3323 TCTCTCTCCCTGTACACACTGGACCCCACTGGCTGAGAAATCTGGGGGTGAGGAGACAA 3382  
Qy 3197 GAAGGAGAGGAAATGTTTCTTGTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 3256  
Db 3383 GAAGGAGAGGAAATGTTTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3442  
Qy 3257 TCTCTCTCCATCAGCTGAAACACTGGACCTGGGGGTAGCCCCGCCAGCCCTCAGTCA 3316  
Db 3443 TCTCTCTCCATCAGCTGAAACACTGGACCTGGGGGTAGCCCCGCCAGCCCTCAGTCA 3502  
Qy 3317 CCCACCTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3376  
Db 3503 CCCACCTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3560  
Qy 3377 AGTAAATATGGGATGGGGGAAGAGGAGCAAGCCAGCCATAGCTTGGGGTGGACA 3436  
Db 3561 AGTAAATATGGGATGGGGGAAGAGGAGCAAGCCAGCCATAGCTTGGGGTGGACA 3619  
Qy 3437 TCTCTAGTGTAGCTGCCACATTTGATTTTCTATAATCAGCTTGGGGTGTGTACATTTTGG 3496  
Db 3620 TCTCTAGTGTAGCTGCCACATTTGATTTTCTATAATCAGCTTGGGGTGTGTACATTTTGG 3679  
Qy 3497 GGGGAGAGACAGATTTTACACTAATATATGAGCCTAGCTTGGAGCAATTTTAAATCC 3556  
Db 3680 GGGGAGAGACAGATTTTACACTAATATATGAGCCTAGCTTGGAGCAATTTTAAAT-CC 3737  
Qy 3557 CTGCACTAGGCAAGTAAATAAGGTTGAGTTTCCAAAAAATAAAAAAAAAAAAAA 3616  
Db 3738 CTGCACTAGGCAAGTAAATAAGGTTGAGTTTCCAAAAAATAAAAAAAAAAAAAA 3797

```

QY 3617 AAAAAA 3622
Db 3798 AAAAAA 3803

RESULT 6
HSTRKE
LOCUS HSTRKE 3554 bp mRNA PRI 23-JAN-1997
DEFINITION H.sapiens TRK E mRNA.
ACCESSION X74979
NID 9400462
VERSION X74979.1 GI:400462
KEYWORDS trk E gene; Trk E protein.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 3554)
AUTHORS Di Marco,E., Cutuli,N., Guerra,L., Cancedda,R. and De Luca,M.
TITLE Molecular cloning of trke, a novel trk-related putative tyrosine
kinase receptor isolated from normal human keratinocytes and widely
expressed by normal human tissues
J. Biol. Chem. 268 (32), 24290-24295 (1993)
JOURNAL 94043265
MEDLINE 2 (bases 1 to 3554)
REFERENCE De Luca,M.
Direct Submission
Submitted (08-SEP-1993); M. De Luca, Istituto Nazionale per la
Ricerca sul Cancro, Lab Differenziamento Cellulare, Viale
Benedetto XV, 10 - 16132 Genova, ITALY
JOURNAL Location/Qualifiers
FEATURES
source
1..3554
/organism="Homo sapiens"
/db_xref="taxon:9606"
/dev_stage="germ line, fetal /adult"
/tissue_type="brain/keratinocytes"
46..2676
/gene="trke"
46..2676
/gene="trke"
/codon_start=1
/product="TrkE"
/protein_id="CAA52915.1"
/db_xref="PID:9400463"
/db_xref="GI:400463"
/translation="MGPEALSLLLLLVAGSDADMKGHFPAKRYALGMODRTIP
SDIASSSSDSTAARSHRLSSDGDGAMCPAGSVFPKEEYLOYDLORLHLVALVGT
QGRHAGLGKEFSRYRLRYSRDGRWGWKDRWGEVLSGNEDEPGVLDLGPVW
ARLYRYPADRWVSLCLVELYGLWLRDGLLSYAPVGTOTMILSEAVYLDNSTYDGH
TVGLQYGLQLADGVGLDDFRKSEQLRVMPGYDYVGWNSHFSFGYVEMEFDFR
LRAFOAMQVHCNNMHTLGLARLPGVECFRRGPAMAWEGEPMRHLNGLNGLDPRARAV
SVPLGRVAFRLQCFRLFAGPMLFSEISFIDVYVNNSSPALGGTFPPAPWPPGPP
TNESSLLEPRGOOPVAKAEGSPATLIGLCVAILLLILLIALLMLRHLRPLLSKA
ERVLEELTVHLSVPGDITLLNNRGPPEPPYQEPPEPRGNPHASPCVPNGSYSG
DYNEPKGAPLLPFPQNSVPHAEADIVTLQGVGGNTYAVPALPGVAGDGPVRY
DFPRSLRERKEKGFQGEVHLCVEVDSFDLPLNVRKHPLLVAVKILRPDA
TKNARDLFKEVIMSKLDPMIIRLLGVCVQDDPLCMITLWENGDLNQFLSAHQLE
DKAAGAGDGGAAOQPTISYPMLLHVAQAIAAGMYRLATLNFVHRDLATRNCLVGEN
FTIKADFQMSRNLYAGDYRVQGRAVLPIRMWAECLIMKGFETTSADVWAGVTLWE
VLMCRAPPGOLTDQVLENAGEFRDGRQVILSRPACFPQGLIYELMLRCWSESE
ORPPFSQLHRFLAEDALNIV"
BASE COUNT 682 a 1064 c 1065 g 743 t
ORIGIN

Query Match 89.2%; Score 3243.6; DB 10; Length 3554;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 3463; Conservative 0; Mismatches 4; Indels 197; Gaps 2;

QY 38 AGAGATGCTGCCCCACCCCTTAGGCGCCGAGGATCAGGAGCTATGGGACCAAGGCC 97
|||||
Db 2 AGAGATGCTGCCCCACCCCTTAGGCGCCGAGGATCAGGAGCTATGGGACCAAGGCC 61

```



```
Db 3191 CTTCTTCTCTCCATCACCTGAACACTGGACCTGGGGTAGCCCCGCCAGCCCTCA 3250
Qy 3312 GTCACCCCACTCCACCTCCAGCTCTGTAGCTAGAACTCTCTAAGCCCTATACGTTTC 3371
Db 3251 GTCACCCCACTCCACCTCCAGCTCTGTAGCTAGAACTCTCTAAGCCCTATACGTTTC 3310
Qy 3372 TGTGAGTAATATTGGGATTTGGGGGAAAGAGGAGGAGCAAGCCCATAGCCCTTGGGGTT 3431
Db 3311 TGTGAGTAATATTGGGATTTGGGGGAAAGAGGAGGAGCAAGCCCATAGCCCTTGGGGTT 3370
Qy 3432 GGACATCTCTAGTGTAGCTGCCCATTCATTTTCTATATCACTTGGGGTTTGTACATT 3491
Db 3371 GGACATCTCTAGTGTAGCTGCCCATTCATTTTCTATATCACTTGGGGTTTGTACATT 3430
Qy 3492 TTTGGGGGAGAGACACAGATTTTACACTAATATATGACCTAGCTAGGCAATTTTA 3551
Db 3431 TTTGGGGGAGAGACACAGATTTTACACTAATATATGACCTAGCTAGGCAATTTTA 3490
Qy 3552 ATCCCTCTCACTAGGAGTAAATAAGGTTTCAGTTTTCACAAAAAATAAAAAA 3611
Db 3491 ATCCCTCTCACTAGGAGTAAATAAGGTTTCAGTTTTCACAAAAAATAAAAAA 3550
Qy 3612 AAAA 3615
Db 3551 AAAA 3554

RESULT 7
HSRETYK1
LOCUS HSRETYK1 3841 bp mRNA PRI 20-APR-1995
DEFINITION H.sapiens EDDR1 gene for receptor tyrosine kinase.
ACCESSION Z29093
NID 9732799
VERSION Z29093.1 GI:732799
KEYWORDS receptor tyrosine kinase.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 3841)
AUTHORS Laval,S., Butler,R., Shelling,A.N., Hanby,A.M., Poulson,R. and
Ganesan,T.S.
TITLE Isolation and characterization of an epithelial-specific receptor
tyrosine kinase from an ovarian cancer cell line
JOURNAL Cell Growth Differ. 5 (11), 1173-1183 (1994)
MEDLINE 95151638
REFERENCE 2 (bases 1 to 3841)
AUTHORS Shelling,A.N., Butler,R., Jones,T., Laval,S., Boyle,J.M. and
Ganesan,T.S.
TITLE Localization of an epithelial-specific receptor kinase (EDDR1) to
chromosome 6q16
JOURNAL Genomics 25 (2), 584-587 (1995)
MEDLINE 95309932
REFERENCE 3 (bases 1 to 3841)
AUTHORS Keding,C.
TITLE Direct Submission
JOURNAL Submitted (17-DEC-1993) Claude Keding, CNRS Laboratoire de
genetique moleculaire-U184, INSERM, 11, rue Humann, Strasbourg,
Alsace, 67085 cedex, FRANCE
FEATURES
source
1. .3841
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RTK6"
/dev_stage="Adult"
/tissue_type="Ovary"
/cell_type="Epithelial cell"
/cell_line="SKOV-3"
/clone_lib="pKS"
sig_peptide 336..423
note="putative signal peptide"
gene 337..2967
```

## CDS

```
/gene="EDDR1"
/337..2967
/standard_name="epithelial discoidin domain receptor"
/codon_start=1
/product="receptor tyrosine kinase"
/protein_id="CAA82335.1"
/db_xref="PIR:G732800"
/db_xref="GI:732800"
/translation="MGPEALSSLLLLLVASGDADMKGHPDKACRVALGMQDRITPD
SDISASSWSSTAAHRSRLSDSGDAMCPAGSVFKEEYIQVDLQRLHLVALVGT
QGRHAGLGKEFSRSLRYSRDRGWRWQKQWQEVISQNEDEGVKLDLGPVW
ARLVRFYPRADRVMSVCLRVLYGLWRDGLLSYTPVQGTMYLSEAVYINDSYDGH
TVGLQYGGGLGOLADGVVGLDDFRKSOELRWPCYDVGWNSHSFSSGYVEMEFEDR
LRAFOAMQVHCNNMHTLGARLPQGVCEFRFRGPMANEGEPHRLNGLDPRARAV
SVPLGGVRARFLQCFRFLAGPWLFSFISIDVNNSSPALGCTFPAPWPPPPPP
TNFSLELEPRGQPPFAAGSPSTAILGLVAIILLIILLIILLIILLIILLIILLI
ERRVEBELTVHLSVPGDTILINRPGPREPPPEPRPNPSPHSPVNGSAYSG
DYMEPEKPGAPLLPQPNSVPHYAEADIVTLQGTGNTYAVPALPPGAVGPGPRV
DFPRSRFLKELGEGQFVEHCEVDSQDLNLSLDFPLNVRKGPLLVAVKILRPDA
TKNARNDLFKEVKIMSRKDPNIIIRLLGVQDDPLCMITDYNHENDLNFLSAHLE
DKAAGAPGQGAAGQPTISYPMLLHVAAQIASCMRYLATLNFVHRLATNCLVGEN
FTIKIADFQMSRNLVAGDIYRVOGRAVLPIRMMAWECILMKFTTASDVWAFGVTLWE
VLMICRAQPFQGLTDEQVIENAGEFFRDOGRQVYLSRPPACPGLYELMLRCWSESE
ORPPFSQLHRELAEDALNTV"
polya_signal 3807
polya_site 3828
BASE COUNT 750 a 1123 c 1126 g 842 t
ORIGIN
Query Match 88.9%; Score 3232.4; DB 10; Length 3841;
Best Local Similarity 94.4%; Pred. No. 0;
Matches 3455; Conservative 0; Mismatches 6; Indels 199; Gaps 2;
Qy 38 AGAGATGTCGCCCCACCCCTTAGGCCGAGGATCAGGAGCTATGGACACAGAGCCC 97
Db 293 AGAGATGTCGCCCCACCCCTTAGGCCGAGGATCAGGAGCTATGGACACAGAGCCC 352
Qy 98 TGTCACTTTTACTGCTGCTGCTTGTGTGCAAGTGGAGATGCTGACATGAAGGACATT 157
Db 353 TGTCACTTTTACTGCTGCTGCTTGTGTGCAAGTGGAGATGCTGACATGAAGGACATT 412
Qy 158 TTGATCTCTCCCAAGTGGCGCTATGCCCTGGGCATCGAGGACCGACCATCCACAGAGTG 217
Db 413 TTGATCTCTCCCAAGTGGCGCTATGCCCTGGGCATCGAGGACCGACCATCCACAGAGTG 472
Qy 218 ACATCTCTGCTTCCAGTCTCTGTCAGATTCACCTGCCGCCGCCACAGAGTTGAGA 277
Db 473 ACATCTCTGCTTCCAGTCTCTGTCAGATTCACCTGCCGCCGCCACAGAGTTGAGA 532
Qy 278 GCAGTGACGGGATGGGCGCTGTCGCCCGAGGTCGTTTCCCAAGGAGGAGGAGT 337
Db 533 GCAGTGACGGGATGGGCGCTGTCGCCCGAGGTCGTTTCCCAAGGAGGAGGAGT 592
Qy 338 ACTTGCAGGTGGATCTAACAGACTGCACCTGTGTGCTGCTGTGGGCGCCACAGGAGCGC 397
Db 593 ACTTGCAGGTGGATCTAACAGACTGCACCTGTGTGCTGCTGTGGGCGCCACAGGAGCGC 652
Qy 398 ATGCCGGGGGCTGGGCAAGAGTTCTCCCGGAGCTACCGGCTGCGTTTACTCCCGGAGT 457
Db 653 ATGCCGGGGGCTGGGCAAGAGTTCTCCCGGAGCTACCGGCTGCGTTTACTCCCGGAGT 712
Qy 458 GTCGCCGCTGGATGGGCTGGAAGGACCGCTGGGGTCAGGAGGTGATCTCAGGCAATGAGG 517
Db 713 GTCGCCGCTGGATGGGCTGGAAGGACCGCTGGGGTCAGGAGGTGATCTCAGGCAATGAGG 772
Qy 518 ACCCTGAGGAGTGGTGTCTGAAGAACCTTGGGCCCCCCTTGGTGGCCGACTGTTTGGCT 577
Db 773 ACCCTGAGGAGTGGTGTCTGAAGAACCTTGGGCCCCCCTTGGTGGCCGACTGTTTGGCT 832
Qy 578 TCTACCCCGGGCTGACCGGGTCTATGAGCTCTGTCTGCGGGTAGAGCTCTATGGCTGCC 637
```











```
Db 2618 GGCAGCTCACCGAGCAGAGCTCATCGAGAACGCGGGGAGTTCTTCCGGGACCGAGGCC 2677
Qy 2678 GGCAGGTGTACTGTCCCGCGCGCTGCCTGCCCGAGGCGCTATATGAGCTGATGCTTC 2737
Db 2678 GGCAGGTGTACTGTCCCGCGCGCTGCCTGCCCGAGGCGCTATATGAGCTGATGCTTC 2737
Qy 2738 GGTCTGAGCGCGGAGTCTGAGCAGCGACACCCCTTTTCCAGCTGCACTCGGTTCTCTGG 2797
Db 2738 GGTCTGAGCGCGGAGTCTGAGCAGCGACACCCCTTTTCCAGCTGCACTCGGTTCTCTGG 2797
Qy 2798 CAGAGGATGCACTCAACACGGTGTGAATCACACATCCAGCTGCCCTCCCTCAGGAGTG 2857
Db 2798 CAGAGGATGCACTCAACACGGTGTGAATCACACATCCAGCTGCCCTCCCTCAGGAGTG 2857
Qy 2858 ATCCAGGGAAGCCAGTGACACTAAACAAGAGGACACAATGGCACCTCTGCCCTTCCCTCC 2917
Db 2858 ATCCAGGGAAGCCAGTGACACTAAACAAGAGGACACAATGGCACCTCTGCCCTTCCCTCC 2917
Qy 2918 TCCGACAGCCCATCACTCTTAATAGAGGAGTGCAGCTGC----- 2959
Db 2918 TCCGACAGCCCATCACTCTTAATAGAGGAGTGCAGCTGC----- 2959
Qy 2959 ----- 2959
Db 2978 GCTGATGCCCTTCTCCCTTCTCTGACACACTCTCATGTCCCTTCTCTTCTCTCTTC 3037
Qy 2959 --AGAAGCCCTGTGCGCCACCCAGCTGGTCTGTGATGGATCCCTCTCCACCTCCCTC 3016
Db 3038 CTAGAGCCCTGTGCGCCACCCAGCTGGTCTGTGATGGATCCCTCTCCACCTCCCTC 3097
Qy 3017 TAGCCATCCCTTTGGGGAAGGTGGGGAATAATAGATAGACACTGGACATGCCCAATT 3076
Db 3098 TAGCCATCCCTTTGGGGAAGGTGGGGAATAATAGATAGACACTGGACATGCCCAATT 3157
Qy 3077 GGAGCAGCTGGCCCTCACTGGACACACTGATTCCTGGAGAGTGGCTGCGCCCTCAGCT 3136
Db 3158 GGAGAACCTGGCCCTCACTGGACACACTGATTCCTGGAGAG-----GAGACCTCAGCT 3211
Qy 3137 TCTCTCTCTCTGTACACACTGGACCCCACTGGCTGAGATCTGGGGGTGAGGAGGACAA 3196
Db 3212 TCTCTCTCTGTACACACTGGACCCCACTGGCTGAGATCTGGGGGTGAGGAGGACAA 3271
Qy 3197 GAAGAGAGGAAATGTTTCTGTGCTGCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 3256
Db 3272 GAAGAGAGGAAATGTTTCTGTGCTGCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 3331
Qy 3257 TCCTCTCCATCACTGAAACACTGGACCTGGGGGTAGCCCGCCCGCCCGCCCTCAGTCA 3316
Db 3332 TCCTCTCCATCACTGAAACACTGGACCTGGGGGTAGCCCGCCCGCCCGCCCTCAGTCA 3391
Qy 3317 CCCCACTTCCCACTGCTGTAGCTAGAACTTCTTAAGCCTATAGCTTCTGTGG 3376
Db 3392 CCCCACTTCCCACTGCTGTAGCTTCTAGCTAGA--CTCTTAAGCCTATAGCTTCTGTGG 3449
Qy 3377 AGTAATAATTCGGATTTGGGGGAAAGAGGAGGACCAAGCCCGCCCGCCCTGAGTGGACA 3436
Db 3450 AGTAATAATTCGGATTTGGGGGAAAGAGGAGGACCAAGCCCGCCCGCCCTGAGTGGACA 3508
Qy 3437 TCTCTAGTGTAGCTGCCACATGATTTTCTATATACATTCAGTGGGTTTGACATTTTGG 3496
Db 3509 TCTCTAGTGTAGCTGCCACATGATTTTCTATATACATTCAGTGGGTTTGACATTTTGG 3568
Qy 3497 GGGGAGAGACACAGATTTTACATAATATATGACCTAGCTGTGAGGCAATTTTAAATCCC 3556
Db 3569 GGGGAGAGACACAGA-TTTTACATAATATATGACCTAGCTGTGAGGCAATTTTAAAT-CC 3626
Qy 3557 CTGCACTAGGAGGTAATAATAAGGTTGAGTTTTCACAAAAAATAAAAAAATAAAAAA 3616
Db 3627 CTGCACTAGGAGGTAATAATAAGGTTGAGTTTTCACAAAAAATAAAAAAATAAAAAA 3686
Qy 3617 AAAAAA 3622
|||||
```

Db 3687 AAAAAA 3692

RESULT 9

AF026259 3659 bp mRNA ROD 27-OCT-1997

LOCUS Mus musculus receptor-like tyrosine kinase (Nep) mRNA, complete cds.

ACCESSION AF026259

VERSION 92564939

KEYWORDS AF026259.1 GI:2564939

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 3659)

AUTHORS Zerlin, M., Julius, M.A. and Goldfarb, M.

TITLE NEP: a novel receptor-like tyrosine kinase expressed in proliferating neuroepithelia

JOURNAL Oncogene 8 (10), 2731-2739 (1993)

MEDLINE 93390947

REFERENCE 2 (bases 1 to 3659)

AUTHORS Zerlin, M., Julius, M.A. and Goldfarb, M.

TITLE Direct Submission

JOURNAL Submitted (22-SEP-1997) Brookdale Center for Developmental and Molecular Biology, Mount Sinai School of Medicine, 1 Gustave Levy Place, New York, NY 10029, USA

FEATURES

Location/Qualifiers

1..3659

/organism="Mus musculus"

/strain="129"

/db\_xref="taxon:10090"

/cell\_type="CCE embryonic stem cells,

differentiation-induced"

1..3659

/gene="Nep"

101..2836

/gene="Nep"

/note="NEP; DDR1"

/codon\_start=1

/product="receptor-like tyrosine kinase"

/protein\_id="AAB81866.1"

/db\_xref="PID:92564940"

/db\_xref="GI:2564940"

/translation="MGTTLSLLLLLTIGDAMKGFHPAKRYALGMODRTIP

DSDISVSSWSDSSTAARSRLESDDGAWCPAGVPKPEEYLVQDLRLHLVVG

TQRHAGLGKEFSRYSRLSRDGRWMDKRWGQEVISGNEGCVVKGDLGPPM

VARLVREYPRADRVMSVCLRVLYGLWRLGLISYTAPOGTWQSLSEVWHLNDSTYD

GYTAGGLQYGLGLADGVVGLDDFRQSELVWPGYDYVWNSQFPTGYVMEFEF

DLRTFTQMVHVCNNMHTLGARLPGGVECRKRGPMAMWEGEPVHALGSLGDPRAR

AISVPLGHHVGRFLQCRFLFAGPWLFESEISFSDVVDSSDTFPPAPWPPGPPPTN

FSSLEPRGQOPVKAEGSPTAILIGCLVAIILLIILALMLRHLWRLLSKADS

RVLSEELTVHLSVPGDILLNNRGPPEPPYQEPPEPRGTPPHSAPCVNGSALLLN

PAYRLLIATYARPPRGPPTPAWKPTNTQACSDYMEPEKPGAPLLPPPPQNSVPH

YAEADIIVTLQVGTGNTYAVPALPGAGDGPVDFPRSLRFKEKLGECQGEVHL

CEVEDPDLYSSFPPISVHKGHPLLVAVKILRPDATKNARDELKEVKIMSKLDNPI

IRLLGVQVDDPLCMITDYMENGDLNOFLSAROLENKATOGLSGDTESDQGPITSYM

LLHVGAIASCMRYLATLNEVHRDLATNCLVGENFTIKIADFGMSRLYAGDYRVQ

GRAVLPTRMNAWECILMGKTTASDWAFGVTLWEVLMCRSQPFQGLTDEQVIRNG

EFFRQDGRQVLSRPPACPTLYELMLRCWSREPEQRPPFAQLHRLADDALNTV"

194..658

/gene="Nep"

/note="encodes discoidin-like domain"

1340..1411

/gene="Nep"

/note="encodes transmembrane domain"

1922..2809

/gene="Nep"

/note="encodes tyrosine kinase domain"

724 a 1095 c 1075 g 765 t

BASE COUNT

ORIGIN

Query Match 65.3%; Score 2375.8; DB 12; Length 3659;  
 Best Local Similarity 81.5%; Pred. No. 0;  
 Matches 2969; Conservative 0; Mismatches 582; Indels 94; Gaps 15;

QY 38 AGAGATGCTGCCCCACCCCTTAGGCCCCGAGGATCAGGAGCTATGGGACAGAGGCC 97  
 DB 57 AGAGATGCTGCCCTCCCTCCATAGACAGAGGATCCAGAGGATGGGACAGAGCCC 116  
 QY 98 TGTCACTTTTACTGTGCTG---CCTTTGGTGGCAAGTGGAGATGCTGACATGAAGGAC 154  
 DB 117 TCTCATCTCTACTGTGCTGCTACTTCTTGGTGCACATTTGGAGATGCTGACATGAAGGAC 176  
 QY 155 ATTTTGAATCCTGCCAAGTGCCTATGCTGCTGGGATGCGAGACCGGACCATCCAGACA 214  
 DB 177 ATTTTGACCTGCCAAGTGCCTATGCTGCTGGGATGCGAGACCGGACCATCTCTGACA 236  
 QY 215 GTGACATCTCTGCTCCAGCTCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 274  
 DB 237 GGTATATCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 296  
 QY 275 AGACAGTGCAGGGATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 334  
 DB 297 AAGCAGTGCAGGATGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 356  
 QY 335 AGTACTTGCAGGTGGATCTACAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 394  
 DB 357 AGTACTTGCAGGTGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 416  
 QY 395 GGCATCCGGGGCTGGGACAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 454  
 DB 417 GCATGCTGGGGTGGGCAAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 476  
 QY 455 ATGGTCCGCTGGATGGGCTGGAGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 514  
 DB 477 ATGGCCCGCTGGATGGGCTGGAGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 536  
 QY 515 AGGACCTGAGGAGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 574  
 DB 537 AGGATCTGGGGAGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 596  
 QY 575 GCTTCTACCCCGCTGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 634  
 DB 597 GCTTCTACCCCGCTGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 656  
 QY 635 GCCTCTGGAGGATGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 694  
 DB 657 GCCTCTGGGGATGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 716  
 QY 695 CTGAGG---CCGCTGCTACCTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 751  
 DB 717 CTGAGGTGATGTATCATCTCAATGATTCACCTTACGATGATATCTGCTGGAGGCTGCT 776  
 QY 752 AGTATGGGGTCTGGGCGAGCTGGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 811  
 DB 777 AGTATGGGGTCTGGGCGAGCTGGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 836  
 QY 812 GTCAGGAGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 871  
 DB 837 GTCAGGAGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 896  
 QY 872 CCAGTGGCTATGAGGATGGATTTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 931  
 DB 897 CCAGGCTACGTGGAGTGGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTG 956  
 QY 932 AGTTCACCTGTACACATGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 991  
 DB 957 AGTTCACCTGTACACATGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1016  
 QY 992 GCTTCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1051  
 DB 1017 GGTTTAAAGGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1076

QY 1052 GCACCTCTGGGGACCCACAGAGCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1111  
 DB 1077 GCAGCCTTGGAGACCCACAGAGCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1136  
 QY 1112 GCTTCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1171  
 DB 1137 GCTTCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1196  
 QY 1172 TCATCTCTGATGTGGTGAACAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1231  
 DB 1197 TCATCTCAGATGTGGTGAAGACTCTCT-----GACACCTTCCACAGAGCC 1244  
 QY 1232 CTG 1291  
 DB 1245 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1304  
 QY 1292 GCCAGAGCCCGTGGGCAAGCCGAGGAGGAGCCGAGCCGAGCCGAGCCGAGCCGAGCCGAG 1351  
 DB 1305 GTCAACAGCCAGTGGGCAAGGAGGAGGAGCCAACTGCCATCTCATTTGGCTGCTGCTG 1364  
 QY 1352 TGGCCATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1411  
 DB 1365 TGGCCATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1424  
 QY 1412 GCAGGCTCTCAGCAAGGCTGAACGAGGAGTGTGGAAGAGGAGCTGACGCTTACCTCT 1471  
 DB 1425 GCGGCTGCTCAGCAAGGCTGACAGCGGCTGTTGGAGGAGGAGCTGACGCTTACCTTT 1484  
 QY 1472 CTG 1531  
 DB 1485 CTG 1544  
 QY 1532 ACCAGAGCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1591  
 DB 1545 ACCAGAGCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1604  
 QY 1592 CTG 1651  
 DB 1605 CTG 1664  
 QY 1652 CTGAGGCGCGGCG 1711  
 DB 1665 CTGAGGCGCGGCG 1724  
 QY 1712 GTGGGACTATATGGAGCTGAGAGCCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1771  
 DB 1725 GTGGGACTATATGGAGCTGAGAGCCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1784  
 QY 1772 ACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1831  
 DB 1785 ACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1844  
 QY 1832 ACACCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1891  
 DB 1845 ACACCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1904  
 QY 1892 ATTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1951  
 DB 1905 ATTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1964  
 QY 1952 TGCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2011  
 DB 1965 TACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2024  
 QY 2012 TGCCTAAGGAGACCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2071  
 DB 2025 TGCCTAAGGAGACCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2084  
 QY 2072 ATGCCAGGATGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2131  
 DB 2085 ATGCCAGGATGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2144  
 QY 2132 TCATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2191

Db 2145 TCATCCGGCTCTGGGTGTGTGTGAGGATGACCCCTCTGCATGATCACAGACTACA 2204  
QY 2192 TGGAGACGGGACCTCAACAGTTCCTCACTGCGCCACAGCTGGAGGACAAAGGACCG 2251  
Db 2205 TGGAGACGGGACCTGACAGCTTCCTCACTGCGCCACAGCTGGAGGACAAAGGACCTC 2264  
QY 2252 AGGGGGCCCCCTGGGACGGGAGGCTCGGACAGGGGCCCCACCATCAGCTACCAATGCTGC 2311  
Db 2265 AGGGGCTCTCTGGGACACAGAGTGTACACAGGGGCCCCACATCAGCTACCTATGCTGT 2324  
QY 2312 TGCATGTGGACAGCCAGATCCGCTCGGCATGCGCTATCTGGCCACACTCAACTTTGTAC 2371  
Db 2325 TACAGTGGGGGCCAGATCCGCTCGGCATGCGCTATCTGGCCACACTCAACTTTGTGC 2384  
QY 2372 ATCGGACCTGGCCACGGGAACTGCCTAGTTGGGGAATAATTCACCATCAAAATGCGAG 2431  
Db 2385 ATCGGACCTGGCCACCGGAACTGCTTGGTGGGGAATAATTCACCATCAAAATGCGCG 2444  
QY 2432 ACTTTGGCATGAGCCGGAACCTCTATGCTGGGGAATTAACCGTGTGACGGGCGGSCAG 2491  
Db 2445 ACTTTGGCATGAGCCGGAATCTACGCTGGGGAATTAACCGTGTGACGGGCGGSCG 2504  
QY 2492 TGCTGCCATCCGCTGGATGCGCTGGAGTGCATCTCATGCGGAAGTTTCAGACTGCGA 2551  
Db 2505 TGCTGCCATCAGGTGGATGCGCTGGAGTGCATCTCATGCGGAAGTTTCACACAGCCA 2564  
QY 2552 GTGAGCTGTGGGCTTTGGTGTGACCTGTGGAGGTGCTGATCCTGTAGGGCCGAGC 2611  
Db 2565 GTGAGCTGTGGGCTTTGGTGTGACCTGTGGAGGTGCTGATCCTGTAGGGTCCGAGC 2624  
QY 2612 CCTTTGGCAGCTCACCGACGAGCAGTGTATCGAGACGCGGGGAGTTCTCCGGGACC 2671  
Db 2625 CCTTTGGCAGCTTACAGATGAGCAGTGTATCGAGATGCGGGAGTTCTTCAGGAGACC 2684  
QY 2672 AGGGCCGAGGTGTACCTGTCCCGGGCGCTGCTGCTGCGGAGGCGCTATATGAGTGA 2731  
Db 2685 AGGGCCGAGGTGTACTGTTCAGGCGGACCGCTGCTGCTGCGGAGGCTGTATGAGTGA 2744  
QY 2732 TGCTTCGCTGTGGAGCGGGAGTGTGAGCAGCGACACCTTTTCCAGCTGCATCGGT 2791  
Db 2745 TGCTTCGCTGTGGAGCGGGAGTGTGAGCAGCGGCGGCTGCTGCGGAGCTTCATCGGT 2804  
QY 2792 TCCTGGCAGAGTGTCACTCAACAGGTGTGAATCAACATCCAGCTGCGCC-CTCCCTCA 2850  
Db 2805 TCCTGGGAGTGTGCTGCTCAACAGGTGTAACTCAGGACCGGAGCCTTTCCCAAT 2864  
QY 2851 GGGAGTGTACAGGGAGCGAGTGTACACTAAACAGAGAGACATGCGACTCTGCC 2910  
Db 2865 GGGAGCATCAGGGGAGC--TGGACTCGAAACAGAGAGAGCTATGGCAGCCGACCC 2922  
QY 2911 CTTCCTCTCCGACAGCGCCATCACCTCTAATAGAG-----GCAGTGAAGTGCAGAAG 2963  
Db 2923 CACCTCTCCATCTTGCCTATCCCAAGAGGAGTGTCTCTGCGGTGGGCTGGGCGCTG 2982  
QY 2964 CCGCTGTGCGCCACCGAGCTGTGCTGTGATGGATGCT-----CTCCA 3008  
Db 2983 CCGAGGAGCAGATACACACTCCCTTCCGTTTCTTCTTCCATCTGCCAAGTGCACCC 3042  
QY 3009 CCGCTCTAGCCATCCCTTGGGAGGGTGGGAGAAATATAGGATAGACATGACAT 3068  
Db 3043 CACCGAGTGGGCTGTGGTGGGATCCTGCTGACTTCTTCAGCTATCCCTTGGGGA 3102  
QY 3069 GCGCATTTGAGCAGCTGGG-----CCCACTGGACAACACTGATTTCTGGA 3115  
Db 3103 GAACGGGGAATATGCTGGGAGTACTGGACAAGGCCCACTGGACACACTGTTCTGGG 3162  
QY 3116 GAGTGGCTGGGCCCCAGCTTCTCTCCCTGTACACACTGGAGCCCACTGGCTGAGA 3175  
Db 3163 AAGGTGAGTGGCCCCAGGTT-----TCTTCTGTGACACACTGGAGCCGCGGCTGAGA 3219  
QY 3176 ATCTG-GGGGTGAGGAGCAAGAGAGGAAATGTTTCTTGTGCTGCTCTCTGTA 3234

Db 3220 ATTAGAGGGGTGAGGGGACAAAGGAGC-----GCCAACCCCTTACA 3259  
QY 3235 CTTGTCTCAGCTTGGGCTTCTTCTCTCCATCACCTGAAACACTGACCTGGGGGTAG 3294  
Db 3260 GCGGTCTCAGCCCTGGCTTCTTCTTCCGCCCTTGACACACTGGACCTGGGGGCGA 3319  
QY 3295 CCCCAGCC-----CAGCCCTCAGTCACCCACCTTCCACCTGCGAGTCTTGTAGCTAGA 3348  
Db 3320 TCCCTGCCCTTGATCCTCAGTCACACACCCCTCTCTCCACCTGCCATGCTACAGCTAGA 3379  
QY 3349 ACTTCTTAAGCCTATAGCTTCTGTGGAGTAATATTTGGGATTTGGGGGAAAGAGGAG 3408  
Db 3380 ACTTCCGAAGCCTCTATGTTTCTGTGGAGAAATATTAG---ATGGGGAACAGAGGAG 3436  
QY 3409 CAACGCCCATAGCCTTGGGTTGGACATCTCTAGCTAGTGCACATTTGATTTTCTA 3468  
Db 3437 CAATAGCTTGAGC--TGGGGTGGGATTTCTATTGAGCTACCAATTTGGTTTCTA 3494  
QY 3469 TAATCAGTTGGGTTTGTACATTTT---GGGGGAGAGACACAGATTTTACACTAATA 3525  
Db 3495 TAATCAGCGGGGTTTGTACATTTTGGGGGAGAGAAACACAGATTTTACACTAATA 3554  
QY 3526 TATGACCTAGCTTGAAGCAATTTAATCCCTGACCTAGCAGGTAATAATAAGTTG 3585  
Db 3555 TATGACCTAGCTTGAAGCGATTTAATCCCTGCTCTAGCAGGTAATAATAAGTTG 3614  
QY 3586 AGTTTTCACAAAAA-----AAAAA-----AAAAA 3630  
Db 3615 AGTTGTTTGTGTTTTTTTTCACAAAAA-----AAAAA 3659

## RESULT 10

## MUSCAX

## LOCUS

## DEFINITION

## ACCESSION

## NID

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## FEATURES

## SOURCE

## 5'UTR

## gene

## CDS

## Location/Qualifiers

## 1. 3674

## /organism="Mus musculus"

## /strain="C57BL"

## /db\_xref="taxon:10090"

## /note="(vector lambda gt10)"

## 1. 106

## /gene="Cak"

## 1. 2842

## /gene="Cak"

## /gene="Cak"

## /codon\_start=1

## /product="receptor kinase"

## /protein\_id="AA05209.1"

## /db\_xref="PIR:G1161063"

## /db\_xref="GI:1161063"

## /translation="MGTGLSSLLLLLVLTIGDAMKGFDPKARYALGMQDRITP

## DSDTSSVSSWSDTAARHSLSSDGDGAMCPAGVPFKEEYLOVDRRLHVALVG

## TOGRHAGLGKESRSLRYLRDRGRWMDKRWGQEVISGNEPDGKVLKDLGPPM

## VARLRYPRADRVMSVCLVELYGLCLWRDGLLSYAPVQGTQMLSEVAVHLDNSYD

## GYTAGGLQYGLGLADGVVLDLDFRQSLRWPGDYVGVNSQSFPTGIVEMEEF

## DLRTFTQMVCNNMHTLGLRPGGVECFKRGKRWAMWEGEPVRLAGSLGDPRAR

## AISVPLGGHVGFRFLQCRFLFAGPMLLLFSEISFISDVVNDSSDTFFPAPWPPGPPPTN



Db 1971 TACACCTGTGTAGTAGAGACCCGGAAGATCTGGTCAGTAGTACTTCCCTATCAGTG 2030.  
Qy 2012 TGGCTAAGGAGACCCCTTTGCTGGTAGCTGTCAAGATCTTTACGGCCAGATGCCACCAAGA 2071  
Db 2031 TGCACAGGAGACCCCTTGTGTGTAGCAGTGAAGATCTCCGGCCAGATGCCACCAAAA 2090  
Qy 2072 ATGCCAGAAATGATTTCTTGAAGAGGTGAAGATCATGTGAGGCTCAAGAGCCCAAAACA 2131  
Db 2091 ATGCCAGAAATGATTTCTTGAAGAGGTGAAGATCATGTGAGGCTCAAGAGCCCAAAACA 2150  
Qy 2132 TCATTGCGCTGCTGGGGTGTGTGAGGAGACGACCCCTCTGCATGATTAAGTACTGACTACA 2191  
Db 2151 TCATTGCGCTGCTGGGGTGTGTGAGGAGATGACCCCTCTGCATGATTAAGTACTGACTACA 2210  
Qy 2192 TGGAGAGCGGAGCTCAACACAGTTCCTCAGTGGCCACAGCTGGAGGACAAGGCGACCG 2251  
Db 2211 TGGAGAGCGGAGTCTGAACACAGTTCCTCAGTGGCCACAGCTGGAGACAAGGCGACCTC 2270  
Qy 2252 AGGGGGCCCTTGGGAGCGGAGCTGGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 2311  
Db 2271 AGGGGGCTCTTGGGAGCACAGAGTCTACCGAGGCGGAGGCGGAGGCGGAGGCGGAGG 2330  
Qy 2312 TGCATGTGGAGCCAGATCGCCTCGGSCATGCGCTATCTGGCCACACTCAACTTTGTAC 2371  
Db 2331 TACAGTGGGGGCGGAGATCGCCTCTGCGATGCGCTTATCTGCGGAGGCGGAGGCGGAGG 2390  
Qy 2372 ATCGGGAGCTGGCCACCGGAACTCGCTAGTTGGGGAATAATTCACCATCAAAATCGCAG 2431  
Db 2391 ATCGGGAGCTGGCCACCGGAACTCGCTAGTTGGGGAATAATTCACCATCAAAATCGCAG 2450  
Qy 2432 ACTTTGGATAGCGGAACTCTATCGTGGGAGTATACCGTGTGAGGCGGAGGCGGAGG 2491  
Db 2451 ACTTTGGATAGCGGAACTCTATCGTGGGAGTATACCGTGTGAGGCGGAGGCGGAGG 2510  
Qy 2492 TGCCTGCCATCGCTGATGCGCTGGAGTGCATCTCATGGGAGGTTTACCGTGTGAGGCGG 2551  
Db 2511 TGCCTGCCATCGCTGATGCGCTGGAGTGCATCTCATGGGAGGTTTACCGTGTGAGGCGG 2570  
Qy 2552 GTGACGTGTGGGCTTTGGTGTGACCTGTGGAGGTGCTGTGCTGTGAGGCGGAGG 2611  
Db 2571 GTGACGTGTGGGCTTTGGTGTGACCTGTGGAGGTGCTGTGCTGTGAGGCGGAGG 2630  
Qy 2612 CTTTGGGAGCTACCGAGCAGCAGTATCAGAGACGCGGGGAGTTCCTCCGGAGC 2671  
Db 2631 CTTTGGGAGCTTACAGATGAGCAGTATCAGAGATGCGGCGAGTTCCTCAGGAGC 2690  
Qy 2672 AGGCGGCGAGTGTACTGTGCGGCGGCTGCTGCTGCGGAGGCGGAGGCGGAGGCGGAG 2731  
Db 2691 AGGCGGCGAGTGTACTGTGCGGCGGCTGCTGCTGCGGAGGCGGAGGCGGAGGCGGAG 2750  
Qy 2732 TGCTTCGGTGTGAGCGGCGGAGTCTGAGCAGCGACACCCCTTTTCCAGCTGCATCGGT 2791  
Db 2751 TGCTTCGGTGTGAGCGGCGGAGTCTGAGCAGCGACACCCCTTTTCCAGCTGCATCGGT 2810  
Qy 2792 TCCTTGGAGAGGATGCACTCAACAGGTGTGAATCAGATCAGCTGCGGCGGAGGCGGAG 2851  
Db 2811 TCCTTGGAGAGGATGCACTCAACAGGTGTGAATCAGATCAGCTGCGGCGGAGGCGGAG 2870  
Qy 2852 GGAGTGTCCAGGGAAGCGAGTGCATCAACTAAACAGAGGAGACAAATGGACCTCTGCC 2911  
Db 2871 GGAAGGCAATCCAGGGAAGCTGACCTGGAACAC - CGAGAGACCTATGCGACCGGACCC 2929  
Qy 2912 TTCCCTCCGAGACGCCATCACTCTAATAG----- 2944  
Db 2930 ACCGTCTCCATACTTGCCCATTTCCCAAGAGGTGAGTGTCTGTCAGGGTTATGGCTGG 2989  
Qy 2944 -----AGGAGTGAAGCTGAGAGGCGGCGGCTGTG----- 2973  
Db 2990 ACTGGGAGTCCGAAGGAATCAGATATACACACTCTTACGTTTCTTCTTCCATCTGCC 3049  
Qy 2973 -----CCACCCAGCTGTGCTGTGGATGGATGCTCTCCACCGCTCTCTAGCCATC 3024

Db 3050 AAGTGCACCCCTCCAGCTGCCCTGTGGTGGGATCCTGCTGACTTCTTCCAGCTATC 3109  
Qy 3025 CCTTGGGAAGGTTGGGAGAAATATAGATAGACACTGGACATGGCCCATTTGGAGCACC 3084  
Db 3110 CCTTGGGAAGAACGGGGCAAAATGCTGGCGTACACTGGACAGG----- 3156  
Qy 3085 TGGGCCCACTGGGACACACACTGATTCCTGGAGAGTGGCTGGCGGCCAGGTTCTCTCTC 3144  
Db 3156 -----CCCACTGGACTTCACTGGTTCTTGGAGGTGACGGTGGCCCGCAGCTT---TCTT 3207  
Qy 3145 CCGTGCACACACTGACCCCACTGGCTGAGATCTTGGGGTGGAGGAGACAGAAGAGA 3204  
Db 3208 CTTGTACACACTGACGCCCGCAGCTGAGAAATTAGAGGGTGGAGGGGACAGAGGAGC- 3267  
Qy 3205 GGAATAATTTTCTTGTGCTGCTCTGCTACTTGTCTCTCAGCTTGGGCTTCTTCTCTC 3264  
Db 3267 -----GCCAACCCCTACAGGGTCTCTCAGCCCTGGCTTCTTCTGTTT 3309  
Qy 3265 CATCACTGAAACACTGACCTGGGGGTAGC--CCGCCCCCAGCCCTCAGTCACCCCCAC 3322  
Db 3310 CGCCCCCTGACACACTGACCTGGGGGTATCCCTGCGTGTATCTCAGTCACACACCCCTCC 3369  
Qy 3323 TTCCCACTGAGCTTCTGAGTAGAATCTCTTAAGCCTATAGCTTCTCTGAGTAAA 3382  
Db 3370 TTCTCACTGCGCACTCTGAGCTAGAACTTCGAAAGCCTCTATGTTCTGTGAGTAAA 3429  
Qy 3383 TATTGGGATTGGGGGAAGAGGAGCAACGCCCATAGCCTTGGGGTTGGACATCTCTA 3442  
Db 3430 TATTAGAT--GGGGAACAGAGGAGCAATAGCTGAGC--TGGGGTGGGGATTCTA 3485  
Qy 3443 GTGTAGCTGCCACATGATTTTCTATTAATCACTTGGGGTTGTACATTTTT---GGGG 3499  
Db 3486 TTGTAGCTACCACATGTTTCTATTAATCACTGCGGGGTTGTACATTTTTTGGGGGA 3545  
Qy 3500 GAGACACACATTTTACACTATATATGACCTAGCTTGGAGCAATTTAATCCCTG 3559  
Db 3546 GAGAACACAGATTTTACACTATATATGACCTAGCTTAAAGCGATTTAATCCCTG 3605  
Qy 3560 CACTAGGAGGTAAATAAAGGTTGAGTT 3589  
Db 3606 TCCTAGGAGGTAAATAAAGGTTGAGTT 3635

RESULT 11  
RATPTK3D LOCUS  
DEFINITION Rattus norvegicus tyrosine kinase receptor (Ptk-3) gene, complete cds.  
ACCESSION L26525  
NID 9432480  
VERSION L26525.1 GI:432480  
KEYWORDS tyrosine kinase receptor.  
SOURCE Rattus norvegicus  
ORGANISM Rattus norvegicus  
REFERENCE 1 (bases 1 to 3743)  
AUTHORS Sanchez,M.P., Tapley,P., Saini,S., He,B., Pulido,D. and Barbacid,M.  
TITLE Multiple tyrosine protein kinases in rat hippocampal neurons: isolation of Ptk-3, a receptor expressed in proliferative zones of the developing brain  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 91 (5), 1819-1823 (1994)  
MEDLINE 94173920  
FEATURES  
Location/Qualifiers  
1..3743  
/organism="Rattus norvegicus"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/dev\_stage="neonate"  
/tissue\_type="brain"  
178..2910  
/gene="Ptk-3"

Query Match	63.6%	Score 2314;	DB 12;	Length 3743;
Best Local Similarity	80.3%;	Pred. No. 0;		
Matches 2971;	Conservative 0;	Mismatches 580;	Indels 149;	Gaps 16;
QY	7	TCGAGTCGACGTTTGACATTGAAGGAATGCCAAGAGATGCTGCCCCACACCCCTTAGGCC 66		
Db	103	TCCTCTCAGCCCTGGGCTCTTCTGCTACCTTGGAGATGCTGCCCTCCCTCCCATAGGCCA 162		
QY	67	GAGGGATCAGGAGCTATGGGACCAAGAGCCCTGTCTATCTTTACTTGCTGTG--CTCTTG 123		
Db	163	GAGGGATCAGGAGCATGGGACAGGACCCCTCTCATCTCTCTGTGTCTGTCTACTCTTG 222		
QY	124	GTGGCAAGTGGAGATGCTGACATGAAGGACATTTTCATCTGCCAAGTCCGCTATGCC 183		
Db	223	GTGACAATTGGAGATGCTGTGATGAAGGGACATTTTGACCCAGCAAGTCCGCTATGCC 282		
QY	184	CTGGGCAATCGAGGACCGGACCATCCACAGACAGTAGTACATCTCTGTCCAGCTCCTGGTCA 243		
Db	283	CTGGGCATCGAGGACCGCACCATTCACAGACAGCATATCTCTGTGCCAGTCTCTGGTCG 342		
QY	244	GATTCACATGCCGCCGCCACAGAGTTTGAGAGACAGTACGGGATGGGGCCTGGTGC 303		
Db	343	GACTCTACGCCGCCGCCACAGAGGCTGGAAGACAGTATGAGATGGGGTCTGGTGC 402		
QY	304	CCCGCAGGTCGGTGTCTTCCCAAGGAGGAGTACTTGCAGGTGGATCTTACAACGACTG 363		
Db	403	CTGTGAGGCGCTGTCTTCCCAAGAGAGGAGTACCTGCAGGTGNCCTTCGTAGGCTA 462		
QY	364	CACCTGGTGGCTCTGGTGGGCACCCAGGACGGCATGCCGGGGCCCTGGGCAAGGAGTTC 423		
Db	463	CACCTGGTGGCTCTGGTGGGCACCCAGGCGGCCCATCTGGGGTCTGGGCAAGGAGTTC 522		



Db 1591 GCGCCAGGCCCCGAGAGCCACCCCTTACAGAAACCCCGGCTCGGGGACTCCAACT 1650  
Qy 1564 CACTCCGCTCTGTCCTCCCAATGGCTCTGCTGCTGCTCTCAATCCAGCTTACCGC 1623  
Db 1651 CACTCTGGCCCTGGTCCCAATGGCTCTGCTGCTCTCAATCCGGCTACCGC 1710  
Qy 1624 CTCTCTTGGCCACTTACGCCCTGCCCCCTGGAGGCCCGGCCCCCAACCCCGCTGG 1683  
Db 1711 CTCCTTCTGGCCACTTACGCCCTGCCCCCTGGAGGCCCGGCCCCCAACCCCGCTGG 1770  
Qy 1684 GCCAAACCCACACACCCAGGCTACAGTGGGACTATATGAGCCCTGAGAACCCAGG 1743  
Db 1771 GCCAAACCCACACACCCAGGCTACAGTGGGACTATATGAGCCCGGAGAACCCGGC 1830  
Qy 1744 GCGCCGCTTCTGCCCCCACTCCCAAGACAGCTGCCCATATGCGGAGGTGACATT 1803  
Db 1831 GCGCCACTTCTACCCCACTCCCAAGACAGCTGCCCATATGCGGAGGTGACATT 1890  
Qy 1804 GTTACCTTSCAGGCGTACCGGGGCAACACCTATGCTGTGCTGCACTGCCCCCAAGG 1863  
Db 1891 GTCACCTTGCAGGCGTACAGGGGCAACACCTACGCTGTTCGCCACTGCCCCAGG 1950  
Qy 1864 GCAGTGGGGATGGCCCCCAGAGTGATTTCCCTCGATCTGCACTCCGCTCAAGGAG 1923  
Db 1951 GCGGTTGGGGATGGCCCCCAGAGTGATTTCCCTCGATCACGGCTCCGTTCAAGGAG 2010  
Qy 1924 AAGCTTGGGAGGCGAGTTTGGGAGGTGCACCTGTGTGAGTGCAGACGCCCTCAAGAT 1983  
Db 2011 AAGCTTGGGAGGCGCAATTTGGGAGGTGATCTGTGAGTGCAGACGCCCTCAAGAT 2070  
Qy 1984 CTGTGCTAGTTGATTTCCCTTAATGTGCGTAAGGACACACCTTTGCTGTGAGTGTG 2043  
Db 2071 CTGTGCTAGTGTCTCCCTATCAGTGTGCAAAAGGACACACCTTTGCTGTGCGAGT 2130  
Qy 2044 AAGATCTTACGGCCAGATGCCACCAAGATGCCAGGATGATTTCTTGAAGAGGTGAAG 2103  
Db 2131 AAGATCTTCCGGCCAGATGCCACCAAAATGCCAGGATGATTTCTTGAAGAGGTGAAG 2190  
Qy 2104 ATCATGTGAGGCTCAAGGACCCAAACATCATTCGGTGTGTCGGCTGTGTGTCAGGAC 2163  
Db 2191 ATCATGTGAGGCTCAAGGACCTTAACATCATCAGGCTCTCGGTGTGTGTGTCAGGAT 2250  
Qy 2164 GACCCCTCTGCATGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2223  
Db 2251 GACCCCTCTGCATGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2310  
Qy 2224 GCGCAGGCTGGAGGACAGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2283  
Db 2311 GCGCAGGCTGGAGGACAGGCTCACTCAGGGGCTTCTGCGGACAGAGATCTGACCGAG 2370  
Qy 2284 GCGCAGGCTGGAGGACAGGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2343  
Db 2371 GCGCAGGCTGGAGGACAGGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2430  
Qy 2344 CGCTATCTGGCCACTCAACTTTGATACATCGGACCTGGCCACCGGAACTGCTAGTT 2403  
Db 2431 CGATATCTGGCCACTCAACTTTGATACATCGGACCTGGCCACCGGAACTGCTAGTT 2490  
Qy 2404 GGGGAAATTTCAACATCAAAATCCAGACTTTGGCATGAGCCGGAACCTTATGCTGGG 2463  
Db 2491 GGGGAAATTTCAACATCAAAATCCAGACTTTGGCATGAGCCGGAATCTCTAGCTGGG 2550  
Qy 2464 GACTATTACCGTGTGAGGGCGGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2523  
Db 2551 GATTATTACCGTGTGAGGGCGGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2610  
Qy 2524 ATCTCATGGGAAGTTCAAGCTGCGAGTGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2583  
Db 2611 ATCTCATGGGAAGTTCAAGCTGCGAGTGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2670  
Qy 2584 GAGTGTCTGATCTGTAGGCCCCAGCCCTTTGGGAGCTCACCGAGGAGGCTATC 2643

Db 2671 GAGTGTCTTATGCTCTGCAAGTTCAGAGCTTCAGAGCCCTTGGGCAACTCAGAGGAGGTATC 2730  
Qy 2644 GAGAACGCGGGAGTTCTTCCGGACACAGGCGCGCAGGTGTACCTGTTCGCCGCGCCT 2703  
Db 2731 GAGAACGCGGGAGTTCTTCCAGGACACAGGCGCGCAGGTATATCTTATCCAGGCCACC 2790  
Qy 2704 GCTTCCGCGAGGCGCTATATGAGCTGATGCTTGGTCTGAGCCGGGAGTCTGAGGAG 2763  
Db 2791 GCTTCCGCGAGGCGCTTATGAGCTGATGCTTCCGCTTGGAGCCGGAGCCGAGGAG 2850  
Qy 2764 CCACACACCTTTTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2823  
Db 2851 CGCGCGCCCTTCCAGGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2910  
Qy 2824 ATACACATCCAGCTGCCC-CTCCCTCAGGAGTGTATCCAGGGAAGCAGTGCACATAA 2882  
Db 2911 ACTAAGGACCTGGCAGCCCTTTTCACTGGAGCCATCCAGGGAAGC--TGGACATAA 2968  
Qy 2883 AACAGAGACACAATGGACCTCTGCCCTTCCCTTCCCTGCGAC----- 2925  
Db 2969 AACAGAGAAAGCTATGCGACCCCGCCACCATCTCCCGCTTCCCGCTTCCCGCTTCCCGCT 3028  
Qy 2925 ----AGCCCATCACCTCTAATAGAGGAGTGCAGCTGCAGAGCCCTCT----- 2971  
Db 3029 CAGTGTCTGAGGCTGGGCTGGGCGCCCGCAGGGAACCGAAGACTCTGTACAGATAC 3086  
Qy 2971 -----CGCCCAACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2992  
Db 3089 GTGCTCCCTTCTGCTTCTTCCCTCCACCTGCCAAGCTTCCACCCAGCTGGCCTGTG 3148  
Qy 2993 GATGGATCTCTCACCCCTCTCTAGCATCCCTTGGGGAAGGTGGGGAAGATATAG 3052  
Db 3149 GTTGGATCTCTCCCTGACTTCTTCCAGCTATCCCTTAGGAAAGACAGGCGGATGTG 3208  
Qy 3053 GATAGACCTGACATGGCCCTTGGAGCAGCTGGGCGCCACTGACACACTGATTTCT 3112  
Db 3209 GGCACACCTGGACAGG-----CCACTGGACAACTGTTCTCT 3249  
Qy 3113 GGAGAGGTGGCTGCGCCCGCCAGCTTCTCTCCCTGTGCACACTGGACCCACTGGCTG 3172  
Db 3250 GGGAGGTGACGCTACCCCGCCAGCT---TCTTCTGTGCACACTGGACCCGCGAG--- 3303  
Qy 3173 AGAATCTGGGGTGGAGGAGCAAGAGGAGGAAATGTTCTTCTGCTGCTGCTGCTGCTGCT 3232  
Db 3303 ---AATGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3343  
Qy 3233 TACTTCTCTGAGCTGGGCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3289  
Db 3344 CAATGCTCTGAGCCCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3403  
Qy 3290 GGTAGCCCC-----GCCCGAGCCCTCAGTCAACCCCGCTTCCCGCTGCGAGTCTTGT 3341  
Db 3404 GGTAGCCCTGCTGTGATCTCTCAGTCAAAACCCCTCTCTCTCTCTCTCTCTCTCTCTCT 3463  
Qy 3342 AGCTAGAACTTC-TCTAGCCTATAGTTCCTGCTGAGTAAATATTTGGATTTGGGGGAA 3400  
Db 3464 CGCTAGAACTTCGCAAAAGCCCTCTGCTTCTCTGCTGAGTAAATATTTAGGAT--GGGAA 3521  
Qy 3401 AGAGGAGCAACGGCCATAGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGG 3460  
Db 3522 AGATGAGCAGTAGCCCGAGGCGGGGGTGGG-ATTCTCTGTAGTACCATTTGG 3580  
Qy 3461 TTTTCTATAACTTACTTGGGGTTTGTACA-TTTTGGGGGAGAGACAGATTTTTTACA 3519  
Db 3581 TTTTCTATAACTTACTTGGGGTTTGTACATTTTGGGGGAGAGACAGATTTTTTACA 3640  
Qy 3520 CTAATATATGACCTAGCTTGGGCAATTTTAAATCCCTGCTAGGAGGTAAATAA 3579  
Db 3641 CTAATATATGACCTAGCTTGGGCAATTTTAAATCCCTGCTAGGAGGTAAATAA 3700  
Qy 3580 AGTTTCTAGTTTCCACAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3619  
Db 3701 AGTTTCTAGTTTCTTGTGTTTTTTTTTCAAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAG 3740



## RESULT 12

180847  
LOCUS 1197 bp DNA PAT 20-MAR-1998  
DEFINITION Sequence 7 from patent US 5709858.  
ACCESSION 180847  
NID 93209137  
VERSION 180847.1 GI:3209137  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE  
AUTHORS Godowski, P.J., Mark, M.R. and Scadden, D.T.  
TITLE Antibodies specific for Rse receptor protein tyrosine kinase  
JOURNAL Patent: US 5709858-A 7 20-JAN-1998;  
FEATURES Location/Qualifiers  
source 1..1197  
BASE COUNT 199 a 351 c 407 g 240 t  
ORIGIN

Query Match 32.9%; Score 1197; DB 5; Length 1197;  
Best Local Similarity 100.0%; Pred. No. 1..le-234;  
Matches 1197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 GATGCTGACATGAAGGACATTTTCATCCTGCCAAGTCCGCTATGCCCTCGGGATGCAG 195  
|||||  
DB 1 GATGCTGACATGAAGGACATTTTCATCCTGCCAAGTCCGCTATGCCCTCGGGATGCAG 60  
QY 196 GACCGGACATCCAGACAGTACATCTCTGCTCCAGCTCTGCTGCTAGATTCACATGCC 255  
|||||  
DB 61 GACCGGACATCCAGACAGTACATCTCTGCTCCAGCTCTGCTGCTAGATTCACATGCC 120  
QY 256 GCCGCCACAGCAGTGTGGAGACAGTACAGGGGATGGGGCTGGTCCCGCAGGGTCG 315  
|||||  
DB 121 GCCGCCACAGCAGTGTGGAGACAGTACAGGGGATGGGGCTGGTCCCGCAGGGTCG 180  
QY 316 GTGTTTCCCAAGGAGGAGTACTTTCAGGTGGATCTACACGACTGCACCTGTGGCT 375  
|||||  
DB 181 GTGTTTCCCAAGGAGGAGTACTTTCAGGTGGATCTACACGACTGCACCTGTGGCT 240  
QY 376 CTGGTGGGACCCAGGAGCGCATCCCGGGGCTGGCAAGGAGTTCTCCCGGAGGTAC 435  
|||||  
DB 241 CTGGTGGGACCCAGGAGCGCATCCCGGGGCTGGCAAGGAGTTCTCCCGGAGGTAC 300  
QY 436 CGGCTGCGTTACTCCGGGATGGTCCCGCTGGATGGCTGGAGGACCGCTGGGGTCAG 495  
|||||  
DB 301 CGGCTGCGTTACTCCGGGATGGTCCCGCTGGATGGCTGGAGGACCGCTGGGGTCAG 360  
QY 496 GAGGTGATCTCAGGCAATGAGGACCTCAGGGAGTGGTGAAGGACCTTGGGCCCC 555  
|||||  
DB 361 GAGGTGATCTCAGGCAATGAGGACCTCAGGGAGTGGTGAAGGACCTTGGGCCCC 420  
QY 556 ATGGTGGCCGACTGGTTCGCTTCTACCCCGGGCTGACCGGGTGCATGAGCGTCTGCTG 615  
|||||  
DB 421 ATGGTGGCCGACTGGTTCGCTTCTACCCCGGGCTGACCGGGTGCATGAGCGTCTGCTG 480  
QY 615 CGGCTAGAGCTATAGCTGCTCTGGAGGATGAGTCTCTGCTTACACCGCCCTGCTG 675  
|||||  
DB 481 CGGCTAGAGCTATAGCTGCTCTGGAGGATGAGTCTCTGCTTACACCGCCCTGCTG 540  
QY 676 GGGCAGACAATCTATTATCTCAGGCGCTGTACCTCAACGACTCCACCTATGACGGACAT 735  
|||||  
DB 541 GGGCAGACAATCTATTATCTCAGGCGCTGTACCTCAACGACTCCACCTATGACGGACAT 600  
QY 736 ACCGTGGGCGGACTCAGTATGGGGTCTGGGCCAGCTGGCAGATGGTGTGGTGGGGCTG 795  
|||||  
DB 601 ACCGTGGGCGGACTCAGTATGGGGTCTGGGCCAGCTGGCAGATGGTGTGGTGGGGCTG 660  
QY 796 GATGACTTTAGGAAGAGTCAGGAGCTGCGGGTCTGGCCAGGCTATGACTATGTGGGATGG 855

DB 661 GATGACTTTAGGAAGAGTCAGGAGCTGCGGGTCTGGCCAGGCTATGACTATGTGGATGG 720  
|||||  
QY 856 AGCAACACACAGCTTCTCCAGTGGCTATGTGGAGATGAGTGTGAGTTTACCGGCTGAGG 915  
|||||  
DB 721 AGCAACACACAGCTTCTCCAGTGGCTATGTGGAGATGAGTGTGAGTTTACCGGCTGAGG 780  
|||||  
QY 916 GCCTTCCAGGCTATGACAGTCCACTGTAAACAACATGCACACGCTGGGAGCCGCTGCTGCT 975  
|||||  
DB 781 GCCTTCCAGGCTATGACAGTCCACTGTAAACAACATGCACACGCTGGGAGCCGCTGCTGCT 840  
|||||  
QY 976 GCGGGGTGGAATGTCTGCTTCCGGCTGCGCTGCCATGCCATGCCATGGAGGGGAGCCCATG 1035  
|||||  
DB 841 GCGGGGTGGAATGTCTGCTTCCGGCTGCGCTGCCATGCCATGGAGGGGAGCCCATG 900  
|||||  
QY 1036 CGCCACAACCTAGGGGCAACCTGGGGGACCCAGAGCCCGGGCTGTCTCAGTGGCCCTT 1095  
|||||  
DB 901 CGCCACAACCTAGGGGCAACCTGGGGGACCCAGAGCCCGGGCTGTCTCAGTGGCCCTT 960  
|||||  
QY 1096 GCGGGCTGTGGCTGCTGCTTCTGCAAGTCCGCTTCTCTTTCGCGGGGCTGGTTACTC 1155  
|||||  
DB 961 GCGGGCTGTGGCTGCTGCTTCTGCAAGTCCGCTTCTCTTTCGCGGGGCTGGTTACTC 1020  
|||||  
QY 1156 TTGAGGGAATCTCTTCTGATGTGGTGAACAATTCCTCTCCGGCACTGGGAGGC 1215  
|||||  
DB 1021 TTGAGGGAATCTCTTCTGATGTGGTGAACAATTCCTCTCCGGCACTGGGAGGC 1080  
|||||  
QY 1216 ACCTTCCCGCAGCCCTTGGTGGCGCTGGCCCACTCCCAACCACTTCCAGCACTTGG 1275  
|||||  
DB 1081 ACCTTCCCGCAGCCCTTGGTGGCGCTGGCCCACTCCCAACCACTTCCAGCACTTGG 1140  
|||||  
QY 1276 GAGCTGGAGCCCAAGAGCCAGCAGCCGCTGGCAAGCCGAGGGAGCCGACCCGCC 1332  
|||||  
DB 1141 GAGCTGGAGCCCAAGAGCCAGCAGCCGCTGGCAAGCCGAGGGAGCCGACCCGCC 1197  
|||||

## RESULT 13

AC004211/c  
LOCUS 48731 bp DNA PRI 17-JUL-1998  
DEFINITION Homo sapiens clone UWG:Y67C112 from 6p21, complete sequence.  
ACCESSION AC004211  
NID 93327405  
VERSION AC004211.1 GI:3327405  
KEYWORDS HTG.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 48731)  
AUTHORS Janer, M.M., Guillaudoux, T., Vu, Q., Kutyavin, T., Harter, H. and Geraghty, D.E.  
TITLE Large scale sequence analysis of the human MHC class I region  
JOURNAL Unpublished (1998)  
REMARK Fred Hutchinson Cancer Research Center  
The Clinical Research Division  
1100 Fairview Ave. N., P.O. Box 19024  
Seattle, WA 98109-1024  
REFERENCE 2 (bases 1 to 48731)  
AUTHORS Geraghty, D.E. and Olson, M.V.  
TITLE Direct Submission  
JOURNAL Submitted (23-FEB-1998) Human Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA  
REFERENCE 3 (bases 1 to 48731)  
AUTHORS Geraghty, D.E. and Olson, M.V.  
TITLE Direct Submission  
JOURNAL Submitted (17-JUL-1998) Human Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA  
REMARK Contact: Daniel E. Geraghty (geraghty@fhcr.org)  
COMMENT On Jul 17, 1998 this sequence version replaced gi:2905849.  
Overlapping Sequences:  
5': UWG:370m23.002  
3': UWG:Y67C092  
-----

## Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp.

Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

Double stranded (DS) coverage: 78.6%

DS or two chemistry coverage: 98.7%

Single stranded regions: 4

## Sequence Validation:

This sequence has been validated by Multiple Complete Digest Mapping. Comparison of the experimentally derived map digest fragments with sequence-predicted fragments is given below. Small fragments with a variable cutoff (approximately 400-600bp) are not mapped and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragment groups are separated by dashed lines.

BgIII	HindIII		Seq	NsiI	
	Map	Seq	Map	Seq	
6380.82	6132.00	1353.33	1317.00	16545.53	16061.00
2157.94	2073.00	21029.50	20255.00	5367.59	5115.00
2205.24	2149.00	1150.03	1118.00	8273.80	7899.00
3050.67	2955.00	9349.57	8800.00	1670.66	1633.00
7318.95	7028.00	12028.04	11426.00	7024.90	6661.00
1742.45	1717.00				
3988.58	3826.00				
6401.18	6122.00				
13937.89	13391.00				

FEATURES  
Source

Location/Qualifiers  
1. 48731  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/map="6p21"  
/clone="CGM1:A146A3"  
/sub\_clone="UWGC:y67c112"  
/cell\_line="CGM1"  
/clone\_lib="Wash U YAC library"  
8  
/note="high quality variation versus 5' overlapping clone"  
832  
/note="high quality variation versus 5' overlapping clone"  
1134  
/note="high quality variation versus 5' overlapping clone"  
1220  
/note="high quality variation versus 5' overlapping clone"  
1497  
/note="high quality variation versus 5' overlapping clone"  
1648  
/note="high quality variation versus 5' overlapping clone"  
2443  
/note="high quality variation versus 5' overlapping clone"  
2583  
/note="high quality variation versus 5' overlapping clone"  
3534  
/note="high quality variation versus 5' overlapping clone"  
3542

variation  
/note="high quality variation versus 5' overlapping clone"  
3715  
/note="high quality variation versus 5' overlapping clone"  
3923  
/note="high quality variation versus 5' overlapping clone"  
3925  
/note="high quality variation versus 5' overlapping clone"  
4447.4798  
/rpt\_family="Alu"  
5432.5703  
/rpt\_family="Alu"  
/rpt\_family="Alu"  
complement(6331.6504)  
/rpt\_family="MIR"  
complement(11000.11096)  
/rpt\_family="MIR"  
13159.13460  
/rpt\_family="Alu"  
15959.16254  
/rpt\_family="Alu"  
16800.16713  
/rpt\_family="Alu"  
complement(16741.17029)  
/rpt\_family="Alu"  
17499.17755  
/rpt\_family="Alu"  
17946.18420  
/rpt\_family="Alu"  
18446.18566  
/rpt\_family="Alu"  
18620.19230  
/rpt\_family="Alu"  
19327.19629  
/rpt\_family="Alu"  
complement(19714.19876)  
/rpt\_family="Alu"  
19991.20271  
/rpt\_family="Alu"  
21408.21698  
/rpt\_family="Alu"  
complement(22064.22343)  
/rpt\_family="Alu"  
complement(22946.23220)  
/rpt\_family="Alu"  
complement(29480.29776)  
/rpt\_family="Alu"  
34016.34143  
/rpt\_family="Alu"  
35865.35989  
/rpt\_family="Alu"  
43351.43620  
/rpt\_family="Alu"  
44069.44316  
/rpt\_family="Alu"  
44359.44637  
/rpt\_family="Alu"  
45421.45550  
/rpt\_family="Alu"  
complement(46394.46937)  
/rpt\_family="Alu"  
complement(47953.48515)  
/rpt\_family="LTR9"  
48882.48731  
/rpt\_family="Alu"  
BASE COUNT 12618 a 13116 c 12767 g 10230 t  
ORIGIN

Query Match 23.08; Score 837.2; DB 11; Length 48731;  
Best Local Similarity 83.7%; Pred.No. 3.1e-161;  
Matches 1063; Conservative 0; Mismatches 3; Indels 204; Gaps 2;  
Oy 2532 GGGGAAGTTCACGACTCGAGCTGACGTGTGGCCCTTGGTGACCTGTGGAGGTCT 2591  
|||||

Db 24866 GGGGAAGTTACGACTGCGAGTACGCTGTGGCCCTTGGTGTGACCCCTGTGGAGGTGCT 24807  
 QY 2592 GATGCTCTGTAGGGCCAGCCCTTTGGGAGCTCACCAGCAGCAGGTTCATCGAGAAGC 2651  
 Db 24806 GATGCTCTGTAGGGCCAGCCCTTTGGGAGCTCACCAGCAGCAGGTTCATCGAGAAGC 24747  
 QY 2652 GGGGAGTTCTTCCGGGACCAAGGCGG----- 2679  
 Db 24746 GGGGAGTTCTTCCGGGACCAAGGCGGAGGTTCAGAGTGGAGAGAGGGAAGATGGGTC 24687  
 QY 2679 ----- 2679  
 Db 24686 CGAGCGGGGACAGAGGCGAGAGTTCTCATCTTGGAGACTAAGAAATATTTGTTCC 24627  
 QY 2679 -----GCAGGTGTACTGTTCGGCGCGGCTGCTCCCGC 2713  
 Db 24626 TGACTCTCATCCACACTGCCAATGAGGTGTACTGTTCGGCGCGGCTGCTCCCGC 24567  
 QY 2714 AGGGCTATATGAGTGTATGCTTCGGTGTGAGCGGAGTCTGAGCAGCAGCACCCT 2773  
 Db 24566 AGGGCTATATGAGTGTATGCTTCGGTGTGAGCGGAGTCTGAGCAGCAGCACCCT 24507  
 QY 2774 TTTCCAGCTGCTATGCTTCTTCCGGAGAGGATGCACTCAACACGGTGTGAATCACACATC 2833  
 Db 24506 TTTCCAGCTGCTATGCTTCTTCCGGAGAGTGCATCAACACGGTGTGAATCACACATC 24447  
 QY 2834 CAGTCCCTCTCCCTCAGGAGTGTATCCAGGGAAGCCAGTGCACACTAAACAAAGAGAC 2893  
 Db 24446 CAGTCCCTCTCCCTCAGGAGAGTGCATCAACACGGTGTGAATCACACATC 24387  
 QY 2894 ACAATGACCTCTGCTCCCTTCCCGACAGCCATCCTTAATAGGCGAGTGA 2953  
 Db 24386 ACAATGACCTCTGCTCCCTTCCCGACAGCCATCCTTAATAGGCGAGTGA 24327  
 QY 2954 ACTG----- 2959  
 Db 24326 ACTGAGGTGGCTGGGCCCCCAGGAGAGTGTATGCCCCCTTCTCCCTTCTGACACACA 24267  
 QY 2959 -----AGAGCCCTCTGCTCCCGCAGCCAGTGTG 2987  
 Db 24266 CTCTATGCTCCCTTCTGCTTCTTCTTCTAGAGCCCTGCTCCCGCAGCCAGTGTG 24207  
 QY 2988 CTGTGATGGATCTCTCCAGCTTCTTCTAGCCATCTTGGGGAAGGTGGGAGAAA 3047  
 Db 24206 CTGTGATGGATCTCTCCAGCTTCTTCTAGCCATCTTGGGGAAGGTGGGAGAAA 24147  
 QY 3048 TATAGATAGACACTGGACATGGCCATTTGGAGCACCTGGGCCCCCAGTGGACAACTGA 3107  
 Db 24146 TATAGATAGACACTGGACATGGCCATTTGGAGCACCTGGGCCCCCAGTGGACAACTGA 24087  
 QY 3108 TTCTGGAGAGTGGCTGGGCCCCCAGTGTCTCTCTCCCTGTACACACTGGACCCACT 3167  
 Db 24086 TTCTGGAGAGTGGCTGGGCCCCCAGTGTCTCTCTCTGTACACACTGGACCCACT 24027  
 QY 3168 GCGTGAATCTGGGGTTCAGAGCAGCAGAGAGAGAGAAATTTCTTGTGCTGTC 3227  
 Db 24026 GCGTGAATCTGGGGTTCAGAGCAGCAGAGAGAGAGAAATTTCTTGTGCTGTC 23967  
 QY 3228 TCTGTACTTGTCTCAGCTTGGGCTTCTCTCTCTCATCCTGAAACACTGGACCTG 3287  
 Db 23966 TCTGTACTTGTCTCAGCTTGGGCTTCTCTCTCTCATCCTGAAACACTGGACCTG 23907  
 QY 3288 GGGGTAGCCCCCAGCCCTCAGTCAACCCCTTCCAGCTGCTTGTAGCTAG 3347  
 Db 23906 GGGGTAGCCCCCAGCCCTCAGTCAACCCCTTCCAGCTGCTTGTAGCTAG 23847  
 QY 3348 AACTTCTTAAGCCCTATAGCTTCTGTGGAGTAAATATTTGGATTTGGGGGAAAGAGGA 3407  
 Db 23846 AACTTCTTAAGCCCTATAGCTTCTGTGGAGTAAATATTTGGATTTGGGGGAAAGAGGA 23787  
 QY 3408 GCAAGGCCCATAGCTTGGGGTTCGACATCTAGTGTAGCTGCCACATTTGATTTTCT 3467  
 Db 23786 GCAAGGCCCATAGCTTGGGGTTCGACATCTAGTGTAGCTGCCACATTTGATTTTCT 23727

QY 3468 ATAATCACTTGGGTTGTACATTTTGGGGAGAGACACAGATTTTACACTAATATA 3527  
 Db 23726 ATAATCACTTGGGTTGTACATTTTGGGGAGAGACACACATTTTACACTAATATA 23667  
 QY 3528 TGGACCTAGCTTGGGCAATTTAAATCCCTGACACTAGGAGGAGTAATAATAAGGTTGAG 3587  
 Db 23666 TGGACCTAGCTTGGGCAATTTAAATCCCTGACACTAGGAGGAGTAATAATAAGGTTGAG 23607  
 QY 3588 TTTTCCACAA 3597  
 Db 23606 TTTTCCACAA 23597

RESULT 14

HSRDK17 1072 bp DNA PRI 22-JUL-1996  
 LOCUS H.sapiens gene encoding discoidin receptor tyrosine kinase, exon 17.  
 DEFINITION

ACCESSION X99034  
 NID g1480254  
 VERSION X99034.1 GI:1480254  
 KEYWORDS discoidin receptor tyrosine kinase.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1072)  
 AUTHORS Playford,M.P., Butler,R.J., Wang,X.C., Katso,R.M., Cooke,I.E. and Ganesan,T.S.  
 TITLE The genomic structure of discoidin receptor tyrosine kinase  
 JOURNAL Genome Res. 6 (7), 620-627 (1996)  
 MEDLINE 96389017  
 REFERENCE 2 (bases 1 to 1072)  
 AUTHORS Ganesan,T.S.  
 TITLE Direct Submission  
 JOURNAL Submitted (30-MAY-1996) T.S. Ganesan, Institute of Molecular Medicine, ICRF Dept of Medical Oncology, 402, Molecular Oncology Laboratories, John Radcliffe Hospital, Headington, Oxford OX3 9DU, UK

FEATURES

source 1..1072  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 misc\_feature 58..198  
 /label=stop  
 exon 58..>1072  
 /label=ex17  
 BASE COUNT 239 a 318 c 264 g 251 t  
 ORIGIN

Query Match 22.7%; Score 826.6; DB 10; Length 1072;  
 Best Local Similarity 91.0%; Pred. No. 4.1e-159;  
 Matches 927; Conservative 0; Mismatches 4; Indels 88; Gaps 1;

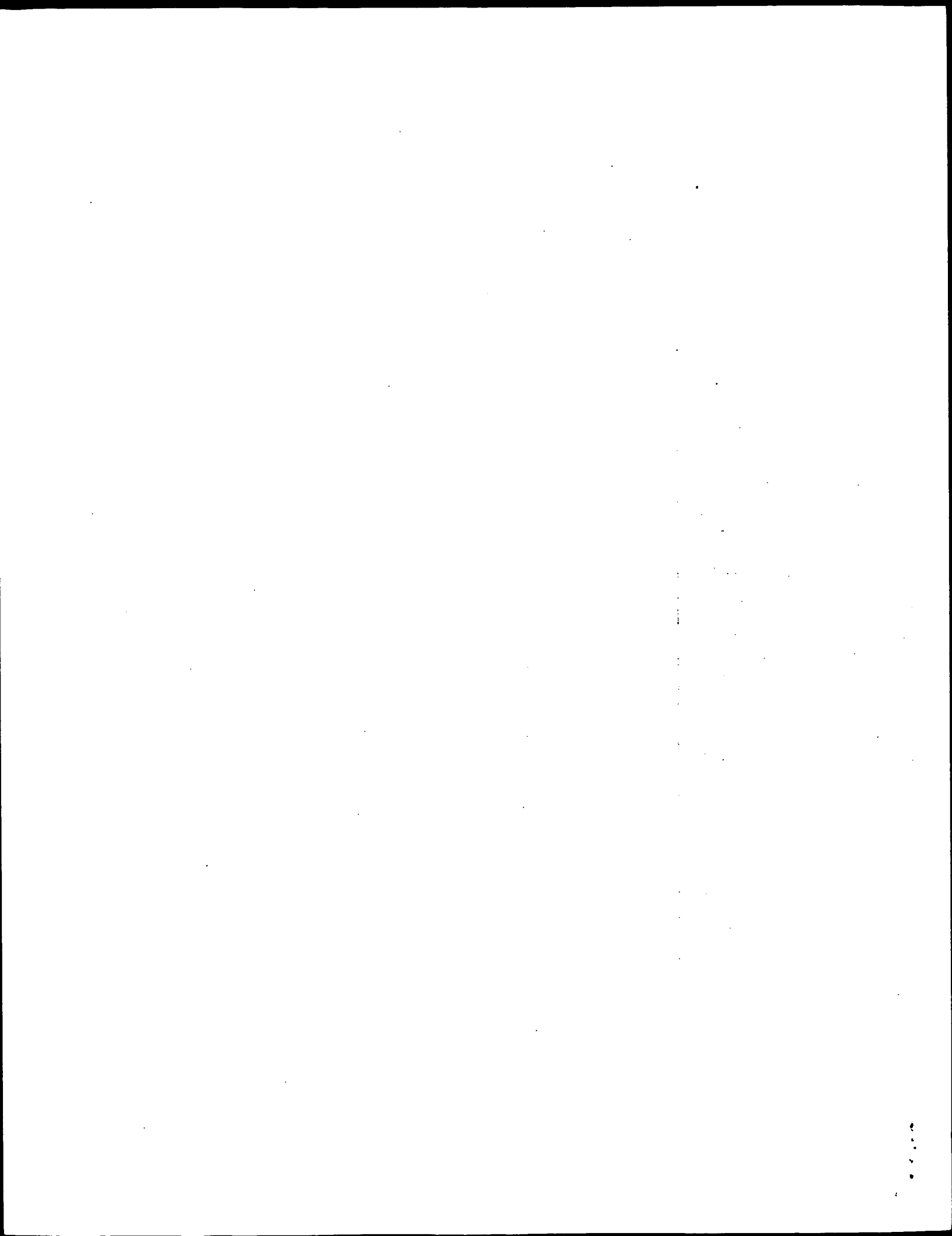
QY 2679 GCAGGTGTACTCTCCCGCGCTGCTGCTGCGGAGCCCTATATGAGCTGATGCTTCG 2738  
 Db 54 GCAGGTGTACTCTGCGCGCGCTGCTGCGGAGCCCTATATGAGCTGATGCTTCG 113  
 QY 2739 GTGCTGGAGCGGAGTCTGAGCAGCGACACCTTTTCCAGCTGCATCGGTTCTCGG 2798  
 Db 114 GTGCTGGAGCGGAGTCTGAGCAGCGACACCTTTTCCAGCTGCATCGGTTCTCGG 173  
 QY 2799 AGAGGATGACTCAACACGGTGTGAATCACATCCAGTGCCTCCCTCAGGAGTGA 2858  
 Db 174 AGAGGATGACTCAACACGGTGTGAATCACATCCAGTGCCTCCCTCAGGAGTGA 233  
 QY 2859 TCAGGGGAGCCAGTGAACACTAAACAGAGGAGACACATGGCACCTTGCCTTCCCT 2918  
 Db 234 TCAGGGGAGCCAGTGAACACTAAACAGAGGAGACACATGGCACCTTGCCTTCCCT 293  
 QY 2919 CCGACAGCCCATCACCTCTAATAGAGGAGTGAAGTGCAGAG----- 2964



QY 2532 GGGGAAGTTACGACTCCGAGTGCAGTGTGGGCTTTGTGTGACCTGTGGAGGTGCT 2591  
|||||  
Db 10746 GGGGAAGTTACGACTCCGAGTGCAGTGTGGGCTTTGTGTGACCTGTGGAGGTGCT 10805  
QY 2592 GATCCTCTGTAGGGCCAGCCCTTTGGGAGCTTCACCGAGGAGCAGGTATCGAGAACGC 2651  
|||||  
Db 10806 GATGCTCTGTAGGGCCAGCCCTTTGGGAGCTTCACCGAGGAGCAGGTATCGAGAACGC 10865  
QY 2652 GGGGAGTTCTTCGGGACACAGGCCGG----- 2680  
|||||  
Db 10866 GGGGAGTTCTTCGGGACACAGGCCGGCAGGTCAGAGTGGAGGAGGGAAGTGGGTC 10925  
QY 2680 ----- 2680  
Db 10926 CGAGNCGGGGAGAGAGGGGCGAGGTTTTCATCTTGGAGACTAAAGAAATAATTGTCC 10985  
QY 2680 -----CAGTGTACCTGTCCCGGCCGCTGCTGCCCGC 2713  
Db 10986 TGACTCTCAACACACTTCCNCAATNCAGGTGTACCTGTCCCGGCCGCTGCTGCCCGC 11045  
QY 2714 AGGSCCTATATGAGCTGATCTTCGGTGTGGAGCCGGGAGTCTGAGCAGCGACACACCT 2773  
Db 11046 AGGSCCTATATGAGCTGATCTTCGGTGTGGAGCCGGGAGTCTGAGCAGCGACACACCT 11105  
QY 2774 TTTCCAGCTGCATCGGTTCTCGCAGAGATGCATCAACACGGTGTGAATCACACATC 2833  
Db 11106 TTTCCAGCTGCATCGGTTCTCGCAGAGATGCATCAACACGGTGTGAATCACACATC 11165  
QY 2834 CAGTGCCTCCCTCAGGAGTGATCCAGGGAGGCCAGTGCACACTAAACAAAGAGGAC 2893  
Db 11166 CAGTGCCTCCCTCAGGAGTGATCCAGGGAGGCCAGTGCACACTAAACAAAGAGGAC 11225  
QY 2894 ACAATGGCCTCTG-CCTTCCCTCCCGACAGCCCATCACCTCTAATAGAGGAGTGA 2952  
Db 11226 ACAATGGCCTCTGCCCTTCCCTCCCGACAGCCCATCACCTCTAATAGAGGAGTGA 11285  
QY 2953 GACTGC----- 2959  
|||||  
Db 11286 GACTGCAGGCTGGGCCACCCAGGGAGTGTATGCCCTTCTCCCTTCTCGGACACATC 11345  
QY 2959 -----AGAGCCCTGTGCCACCCAGCTGTCTCTG 2990  
Db 11346 TCATGTCCCTTCTCTTCTTCTAGAGCCCTGTGCCACCCAGCTGTCTCTG 11405  
QY 2991 TGGATGGATCTCTCCACCTCTCTAGCCATPCCCTTGGGAAGGGTGGGAGAAATAT 3050  
Db 11406 TGGATGGATCTCTCCACCTCTAGCCATPCCCTTGGGAAGGGTGGGAGAAATAT 11465  
QY 3051 AGGATAGACACTGGACATGGCCCATTTGGAGCACCTGGGCCCCACTGGACAACTGATTC 3110  
Db 11466 AGGATAGACACTGGACATGGCCCATTTGGAGCACCTGGGCCCCACTGGACAACTGATTC 11525  
QY 3111 CTGGAGAGTGGCTGGGCCCCAGCTTCTCTCCCTGTACACACTGGACCCCACTGGC 3170  
Db 11526 CTGGAGAGTGGCTGGGCCCCAGCTTCTCTCCCTGTACACACTGGACCCCACTGGC 11585  
QY 3171 TGAGAACTCTGGGGTGGAGGACAGAGAGGAGAAATGTTCCCTGTGCTGTCTCC 3230  
Db 11586 TGAGAACTCTGGGGTGGAGGACAGAGAGGAGAGAAATGTTCCCTGTGCTGTCTCC 11645  
QY 3231 TGTACTTGTCTCAGCTTGGGCTTCTTCTCTCCATCACCTGAAACACTGGACCTGGGG 3290  
Db 11646 TGTACTTGTCTCAGCTTGGGCTTCTTCTCTCCATCACCTGAAACACTGGACCTGGGG 11705  
QY 3291 GTAGCCCCCGCCAGCCCTCAGTCA-CCCCCACTTCCCACTGCACTTGTAGTAGAA 3349  
Db 11706 GTAGCCCCCGCCAGCCCTCAGTCACTCCCCCTTCCCACTTGTAGTAGAA 11765  
QY 3350 CTTCTTAAGCCTATACGTTTCTGTGGAGTAAATTTGGGAGGAGGAGGAGC 3409  
Db 11766 CTTCTTAAGCCTATACGTTTCTGTGGAGTAAATTTGGGAGGAGGAGGAGC 11825

QY 3410 AACGCCCATAGCCTTGGGTTGGACATCTCTAGCTAGCTGCCACATTTGATTTTCTAT 3469  
Db 11826 AACGCCCATAGCCTTGGGTTGGACATCTCTAGCTAGCTGCCACATTTGATTTTCTAT 11885  
QY 3470 AATCAGTTGGGTTGTACATTTTGGGGGAGAGACACAGATTTTACACTAATATG 3529  
Db 11886 AATCAGTT-GGTTTGTACATTTTGGGGGAGAGACACAGATTTTACACTAATATG 11944  
QY 3530 GACCTAGCTTCAGGCAATTTTAAATCCCTGCACCTAGCAGGTAATAATAAGGTTGAGTT 3589  
Db 11945 GACCTAGCTTCAGGCAATTTTAAATCCCTGCACCTAGCAGGTAATAATAAGGTTGAGTT 12004  
QY 3590 TTCCAC 3595  
Db 12005 TTCCAC 12010

Search completed: November 4, 1999, 07:31:19  
Job time: 6478 sec



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 1999, 23:46:05 ; Search time 19.79 Seconds  
(without alignments)  
1848.398 Million cell updates/sec

Title: US-08-170-558-4  
Perfect score: 4903  
Sequence: 1 MGPEALSLLLLLLVASGDA.....QRPFSQLHRLAEDALNTV 913  
Scoring table: BLOSUM62  
Searched: 129810 seqs, 40065486 residues  
Database: PIR\_60.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	4889	99.7	913	2	A48280	receptor tyrosine
2	4664.5	95.1	876	2	A49508	protein-tyrosine k
3	4584.5	93.5	910	2	A53137	tyrosine kinase re
4	2415	49.3	855	2	S42621	protein-tyrosine k
5	2352	48.0	819	2	I48859	tyro 10 receptor k
6	1154	23.5	220	2	I80329	receptor tyrosine
7	952	19.4	248	2	A4547	receptor tyrosine
8	924.5	18.9	183	2	I59442	receptor tyrosine
9	661	13.5	818	1	S44098	brain-derived neur
10	659.5	13.5	821	1	S06943	brain-derived neur
11	659.5	13.5	821	1	A39667	brain-derived neur
12	658	13.4	822	1	A56853	brain-derived neur
13	653.5	13.3	799	1	TVRTTB	nerve growth facto
14	651.5	13.3	825	2	A55178	neurotrophin recep
15	642.5	13.1	825	1	A40026	neurotrophin-3 rec
16	639	13.0	790	1	TVHUTT	nerve growth facto
17	633.5	12.9	839	1	I73632	neurotrophin-3 rec
18	628	12.8	803	4	S23741	hypothetical TPR/T
19	623.5	12.7	803	1	S35695	neurotrophin-3 rec
20	621	12.7	282	2	I51236	brain-derived neur
21	614	12.5	852	2	I51259	tyrosine kinase C
22	608.5	12.4	946	2	A47299	ror-related recept
23	603	12.3	871	2	I48696	gene Nsk2 protein
24	601.5	12.3	881	2	I48697	protein-tyrosine k
25	601.5	12.3	605	2	S67815	protein-tyrosine k
26	600.5	12.2	883	2	S60738	protein-tyrosine k
27	599	12.2	873	2	S60740	protein-tyrosine k
28	591.5	12.1	685	2	A48289	neurotrophic recep
29	580	11.8	981	1	F0MVMG	gag-abl polyprotei
30	546	11.1	937	2	A45082	neurotrophic recep
31	544.5	11.1	402	1	TVFVUR	protein-tyrosine k
32	534.5	10.9	1382	1	INHUR	insulin receptor p
33	534.5	10.9	2338	2	I73957	c-ros-1 tyrosine k
34	534.5	10.9	2317	2	I56752	c-ros-1 tyrosine k
35	534	10.9	2347	1	TVHURS	kinase-related pro
36	533.5	10.9	2311	1	TVCHSR	kinase-related pro
37	530.5	10.8	1372	2	A34157	insulin receptor p
38	530	10.8	2340	2	I48310	kinase-related pro
39	528	10.8	1520	1	TVFFA	protein-tyrosine k

ALIGNMENTS

RESULT 1  
A48280  
receptor tyrosine kinase - human  
C:Species: Homo sapiens (man)  
C>Date: 31-May-1996 #sequence\_revision 31-May-1996 #text\_change 16-Dec-1998  
C:Accession: A48280  
R:Johnson, J.D.; Edman, J.C.; Rutter, W.J.  
Proc. Natl. Acad. Sci. U.S.A. 90, 5677-5681, 1993  
A>Title: A receptor tyrosine kinase found in breast carcinoma cells has an extracellu  
A:Reference number: A48280; MUID:93296201  
A:Accession: A48280  
A>Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1-913 <RES>  
A:Cross-references: GB:L11315; NID:g403386; PID:g403387  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; discoidin I amino-  
F:30-185/Domain: discoidin I amino-terminal homology <DNI>  
F:608-912/Domain: protein kinase homology <KIN>  
F:616-624/Region: protein kinase ATP-binding motif

Query Match	99.7%	Score 4889	DB 2	Length 913
Best Local Similarity	99.7%	Pred. No. 1.5e-263		
Matches	910	Conservative	2	Mismatches 1; Indels 0; Gaps 0;
Qy	1	MGPEALSLLLLLLVASGDADMGHFDPAKRYALGMQDRTPDSDISASSWSDDTAAR	60	
Db	1	MGPEALSLLLLLLVASGDADMGHFDPAKRYALGMQDRTPDSDISASSWSDDTAAR	60	
Qy	61	HSRLSSDGDGAWCPAGSVFPEEYQLQDLRLHLVALVGTGRHAGGLGKFSRYRL	120	
Db	61	HSRLSSDGDGAWCPAGSVFPEEYQLQDLRLHLVALVGTGRHAGGLGKFSRYRL	120	
Qy	121	RYSDGRRWGKDWGEVIGSNEDEGEGVVKDLGPPMVARLVRFPRADRVMSVCLRV	180	
Db	121	RYSDGRRWGKDWGEVIGSNEDEGEGVVKDLGPPMVARLVRFPRADRVMSVCLRV	180	
Qy	181	ELYGCLWRDGLLSYTAAPVQGTMYLSEAVYLNDSYDGTGGLQYGGGLGADGVGLDD	240	
Db	181	ELYGCLWRDGLLSYTAAPVQGTMYLSEAVYLNDSYDGTGGLQYGGGLGADGVGLDD	240	
Qy	241	FRKSOELRVWPGYDYGVGNSHSFSSGVYEMEFEDRLRAFOAMOVHNCNNHTLGLARLPGG	300	
Db	241	FRKSOELRVWPGYDYGVGNSHSFSSGVYEMEFEDRLRAFOAMOVHNCNNHTLGLARLPGG	300	
Qy	301	VECFRRGPAMAWGEPEPRHNLGGLGDPARAVSVPLGGVAVFLQCRFLFAGPWLIFS	360	
Db	301	VECFRRGPAMAWGEPEPRHNLGGLGDPARAVSVPLGGVAVFLQCRFLFAGPWLIFS	360	
Qy	361	EISFSDVYVNNSSPALGGTFFPAPWMPGGPPTNFSSLELEPRGQQPVAKPEGSPTAILI	420	
Db	361	EISFSDVYVNNSSPALGGTFFPAPWMPGGPPTNFSSLELEPRGQQPVAKPEGSPTAILI	420	
Qy	421	GCLVAIIILLIITALLMLRWLHWRLLSKAERRVLEELTVHLSVPGDTILINRPGPRE	480	
Db	421	GCLVAIIILLIITALLMLRWLHWRLLSKAERRVLEELTVHLSVPGDTILINRPGPRE	480	
Qy	481	PPYQEPGRGNPHSPACVPGNSALLSNPAYRLLLATYARPPRPGPPTPAWAKPTNT	540	
Db	481	PPYQEPGRGNPHSPACVPGNSALLSNPAYRLLLATYARPPRPGPPTPAWAKPTNT	540	
Qy	541	QAYSGDYMEPEKPGAPLLPPPPQNSVPHYAEADIVTLOGVTGGNTYAVPALPPGAVGDGP	600	

Db 541 QAYSGDYNEPEKPGAPLLPPPPQNSVPHYAEADIVTLOGVTGGNTYAVPALPPGAVGDP 600  
QY 601 PRVDPPRSRLREKEKLGEGQGEVHLCVDSQDVLVSLDFPLNVRKGHPLLVAVKILRPD 660  
Db 601 PRVDPPRSRLREKEKLGEGQGEVHLCVDSQDVLVSLDFPLNVRKGHPLLVAVKILRPD 660  
QY 661 ATKARNDFLEKVKTMSRLKDPNIIRLLGVCVQDDPLCMITDYMGDNLQFSAHQLED 720  
Db 661 ATKARNDFLEKVKTMSRLKDPNIIRLLGVCVQDDPLCMITDYMGDNLQFSAHQLED 720  
QY 721 KAAEGAPGDGQAAGPTISYPMLLHVAQAISGMRYLATLNFVHRDLATRNCLVGENFTI 780  
Db 721 KAAEGAPGDGQAAGPTISYPMLLHVAQAISGMRYLATLNFVHRDLATRNCLVGENFTI 780  
QY 781 KIADFGMSRLNLYAGDYRYVQGRAVLPIRMAWECILMGKFTTASDVWAFGYTLWEVLMCL 840  
Db 781 KIADFGMSRLNLYAGDYRYVQGRAVLPIRMAWECILMGKFTTASDVWAFGYTLWEVLMCL 840  
QY 841 RAQPFQGLTDEQVIENAGEFFRDQGRQVYLSRPPACPOGLYELMLRCWSRESEQRPPFSQ 900  
Db 841 RAQPFQGLTDEQVIENAGEFFRDQGRQVYLSRPPACPOGLYELMLRCWSRESEQRPPFSQ 900  
QY 901 LHRFLAEDALNTV 913  
Db 901 LHRFLAEDALNTV 913

RESULT 2  
A49508  
protein-tyrosine kinase (EC 2.7.1.112) trkE precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 10-Nov-1995 #sequence\_revision 10-Nov-1995 #text\_change 16-Dec-1998  
C:Accession: A49508; I38358; S37402  
R:Di Marco, E.; Cutuli, N.; Guerra, L.; Cancedda, R.; De Luca, M.  
J. Biol. Chem. 268, 24290-24295, 1993  
A:Title: Molecular cloning of trkE, a novel trk-related putative tyrosine kinase receptor  
A:Reference number: A49508; MUID:94043265  
A:Accession: A49508  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-876 <DIA>  
A:Cross-references: EMBL:X74979; NID:9400462; PID:9400463  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; discoidin I amino-ter  
C:Keywords: ATP; phosphotransferase  
F:30-185/Domain: discoidin I amino-terminal homology <DNI>  
F:571-875/Domain: protein kinase homology <KIN>  
F:579-587/Region: protein kinase ATP-binding motif

Query Match 95.18; Score 4664.5; DB 2; Length 876;  
Best Local Similarity 95.8%; Pred. No. 3.8e-251;  
Matches 875; Conservative 0; Mismatches 1; Indels 37; Gaps 1;

QY 1 MGPEALSLLLLLVASGDADMKGHFPAKCRYALGMQDRTPDSDISASSSSDSTAAR 60  
Db 1 MGPEALSLLLLLVASGDADMKGHFPAKCRYALGMQDRTPDSDISASSSSDSTAAR 60  
QY 61 HSRLESSDGDGAWCPAGSVFPKEEYLQVDLQRLHLVALVGTQGRHAGGLGKEFSRYRL 120  
Db 61 HSRLESSDGDGAWCPAGSVFPKEEYLQVDLQRLHLVALVGTQGRHAGGLGKEFSRYRL 120  
QY 121 RYSDRGWRWGWKDRGQGEVLSGNEDEGCVVKGDPMPVRLVRFYPRADRVMSVCLRV 180  
Db 121 RYSDRGWRWGWKDRGQGEVLSGNEDEGCVVKGDPMPVRLVRFYPRADRVMSVCLRV 180  
QY 181 ELYGCLWRDGLLSYAPVQGMYSLEAVYLNDSYDGTGGLQYGLGOLADGVGLDD 240  
Db 181 ELYGCLWRDGLLSYAPVQGMYSLEAVYLNDSYDGTGGLQYGLGOLADGVGLDD 240  
QY 241 FRKSOELRVPGDYVGVNSHNSFSGVEFEFDRLRAFAQMVCHNNHMTLGARLPGG 300  
Db 241 FRKSOELRVPGDYVGVNSHNSFSGVEFEFDRLRAFAQMVCHNNHMTLGARLPGG 300

QY 301 VECRRFRGAMAWEGEPMRHNLGGNLGDPARAVSVPLGGRVAREFLQCRFLFAGPWLLFS 360  
Db 301 VECRRFRGAMAWEGEPMRHNLGGNLGDPARAVSVPLGGRVAREFLQCRFLFAGPWLLFS 360  
QY 361 ETSFISDVYNNSSPALGGTFPPAPWPPGPPPTNFSSLELEPRGOOPVAKPGSPTAILI 420  
Db 361 ETSFISDVYNNSSPALGGTFPPAPWPPGPPPTNFSSLELEPRGOOPVAKPGSPTAILI 420  
QY 421 GCUVAIILLLLIIALMLWRLHRRLLSKAERRVLEELTVHLSVPGDTILINNPGPRE 480  
Db 421 GCUVAIILLLLIIALMLWRLHRRLLSKAERRVLEELTVHLSVPGDTILINNPGPRE 480  
QY 481 PPYQEPFRGPNPPHSAPCPVNGSALLLSNPAYRLLLATYARPPRGPGPTPAWAKPTNT 540  
Db 481 PPYQEPFRGPNPPHSAPCPVNGS----- 505  
QY 541 QAYSGDYNEPEKPGAPLLPPPPQNSVPHYAEADIVTLOGVTGGNTYAVPALPPGAVGDP 600  
Db 505 -AYSGDYNEPEKPGAPLLPPPPQNSVPHYAEADIVTLOGVTGGNTYAVPALPPGAVGDP 563  
QY 601 PRVDPPRSRLREKEKLGEGQGEVHLCVDSQDVLVSLDFPLNVRKGHPLLVAVKILRPD 660  
Db 564 PRVDPPRSRLREKEKLGEGQGEVHLCVDSQDVLVSLDFPLNVRKGHPLLVAVKILRPD 623  
QY 661 ATKARNDFLEKVKTMSRLKDPNIIRLLGVCVQDDPLCMITDYMGDNLQFSAHQLED 720  
Db 624 ATKARNDFLEKVKTMSRLKDPNIIRLLGVCVQDDPLCMITDYMGDNLQFSAHQLED 683  
QY 721 KAAEGAPGDGQAAGPTISYPMLLHVAQAISGMRYLATLNFVHRDLATRNCLVGENFTI 780  
Db 684 KAAEGAPGDGQAAGPTISYPMLLHVAQAISGMRYLATLNFVHRDLATRNCLVGENFTI 743  
QY 781 KIADFGMSRLNLYAGDYRYVQGRAVLPIRMAWECILMGKFTTASDVWAFGYTLWEVLMCL 840  
Db 744 KIADFGMSRLNLYAGDYRYVQGRAVLPIRMAWECILMGKFTTASDVWAFGYTLWEVLMCL 803  
QY 841 RAQPFQGLTDEQVIENAGEFFRDQGRQVYLSRPPACPOGLYELMLRCWSRESEQRPPFSQ 900  
Db 804 RAQPFQGLTDEQVIENAGEFFRDQGRQVYLSRPPACPOGLYELMLRCWSRESEQRPPFSQ 863  
QY 901 LHRFLAEDALNTV 913  
Db 864 LHRFLAEDALNTV 876

RESULT 3  
A53137  
Tyrosine kinase receptor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 16-Dec-1998  
C:Accession: A53137  
R:Sanchez, M.P.; Tapley, P.; Saini, S.S.; He, B.; Pulido, D.; Barbacid, M.  
Proc. Natl. Acad. Sci. U.S.A. 91, 1819-1823, 1994  
A:Title: Multiple tyrosine protein kinases in rat hippocampal neurons: isolation of p  
A:Reference number: A53137; MUID:94173920  
A:Accession: A53137  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-910 <RES>  
A:Cross-references: GB:L26525; NID:9432480; PID:9432481  
C:Genetics:  
A:Gene: Ptk-3  
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; discoidin I amino  
F:31-186/Domain: discoidin I amino-terminal homology <DNI>  
F:605-905/Domain: protein kinase homology <KIN>  
F:613-621/Region: protein kinase ATP-binding motif

Query Match 93.5%; Score 4584.5; DB 2; Length 910;  
Best Local Similarity 93.5%; Pred. No. 1.1e-246;  
Matches 855; Conservative 20; Mismatches 34; Indels 5; Gaps 2;



QY 1 MGPEALSS-LLLLLIVASGDADMKGHFDPKACRYALGMDRTIPDSISASSWSDDSTAA 59  
 Db 1 MGPTUSSLLLLLLLLITIGDADMKGHFDPKACRYALGMDRTIPDSISASSWSDDSTAA 60  
 QY 60 RHRSLSSDGDGAWCPAGSPVPEKEEYLOVDLQRLHLVALVGTQGRHAGGLGKEFSRYSR 119  
 Db 61 RHRSLSSDGDGAWCPAGSPVPEKEEYLOVDLQRLHLVALVGTQGRHAGGLGKEFSRYSR 120  
 QY 120 LRYSRDGRWGWKDRWGQEVISGNEDEPGVVLKDLGPPMVARLYRFPYPRADRVMSVCLR 179  
 Db 121 LRYSRDGRWGWKDRWGQEVISGNEDEPGVVLKDLGPPMVARLYRFPYPRADRVMSVCLR 180  
 QY 180 VELYGCLWRDGLLSYAPVGTQMTYSEAYLYNDSTYDGHVTCGLQYGLGLQADGVGLD 239  
 Db 181 VELYGCLWRDGLLSYAPVGTQMTYSEAYLYNDSTYDGHVTCGLQYGLGLQADGVGLD 240  
 QY 240 DFKSKOELRWPCYDYGWNSHFSFGYVEMEFEDRLRAFQAMQVHCNNHHTILGARLPG 299  
 Db 241 DFKSKOELRWPCYDYGWNSHFSFGYVEMEFEDRLRAFQAMQVHCNNHHTILGARLPG 300  
 QY 300 GVECFRRGPNAMWEGEPHRLHGNLGDPRARAVSVPLGGVRARFLQCRFLFAGPWLFF 359  
 Db 301 GVECFRRGPNAMWEGEPHRLHGNLGDPRARAVSVPLGGVRARFLQCRFLFAGPWLFF 360  
 QY 360 SEISFISDVVNSSPALGTFEPAPWPPGPPPTNFSSLELEPRCQOQVAKPEGSPTAIL 419  
 Db 361 SEISFISDVVNSSPALGTFEPAPWPPGPPPTNFSSLELEPRCQOQVAKPEGSPTAIL 416  
 QY 420 IGLVAIILLLLLIIALMLRHLWRLLSKAERRVLEELTVHLSVPGDTILNNRPGPR 479  
 Db 417 IGLVAIILLLLLIIALMLRHLWRLLSKAERRVLEELTVHLSVPGDTILNNRPGPR 476  
 QY 480 EPPVQEPFRGNPNPHSPAPVNGSALLSNPAYRLLATYARPPGPPPTPAWAKPTN 539  
 Db 477 EPPVQEPFRGNPNPHSPAPVNGSALLSNPAYRLLATYARPPGPPPTPAWAKPTN 536  
 QY 540 TOAYSGDYMEPEKPGAPILLPPPPQNSVPHYADIVTLQGTGNTYAVPALPGAVGDG 599  
 Db 537 TOAYSGDYMEPEKPGAPILLPPPPQNSVPHYADIVTLQGTGNTYAVPALPGAVGDG 596  
 QY 600 PPRVDFPRSLRFLKELGEGQFGEVHLCVDSQDLVSLDFPLNVRKGLHLLVAVKILRP 659  
 Db 597 PPRVDFPRSLRFLKELGEGQFGEVHLCVDSQDLVSLDFPLNVRKGLHLLVAVKILRP 656  
 QY 660 DATKARNDFLKEVKIMSLRDKPNIIRLLGVCVQDDPLCMITDYMGNDLNOFTLSAHOLE 719  
 Db 657 DATKARNDFLKEVKIMSLRDKPNIIRLLGVCVQDDPLCMITDYMGNDLNOFTLSAHOLE 716  
 QY 720 DKAAEGAPGDGQAAGQPTISYPMILHLVAAQIASGMRYLATLNFVHRDLATRNCLVGENFT 779  
 Db 717 NKVTOGLPDRSDGPTISYPMILHLVAAQIASGMRYLATLNFVHRDLATRNCLVGENFT 776  
 QY 780 IKIADFGMSRLNYAGDYRVQGRAVLPPIRMWAWECILMKFTTASDVWAFGVTILWEVIML 839  
 Db 777 IKIADFGMSRLNYAGDYRVQGRAVLPPIRMWAWECILMKFTTASDVWAFGVTILWEVIML 836  
 QY 840 CRAQPFQGLTDEQVTEENAGEFRDQGVYLSRPPACQGLYELMLRCWSRESQRPFFS 899  
 Db 837 CRAQPFQGLTDEQVTEENAGEFRDQGVYLSRPPACQGLYELMLRCWSRESQRPFFS 896  
 QY 900 QLHRELAEDALNTV 913  
 Db 897 QLHRELAEDALNTV 910

RESULT 4  
 S42621  
 protein-tyrosine kinase (EC 2.7.1.112) - human  
 C:Species: Homo sapiens (man)  
 C:Date: 07-Sep-1994 #sequence\_revision 10-Nov-1995 #text\_change 17-Mar-1999  
 R:karn, T.; Holtrich, U.; Braeuning, A.; Boelme, B.; Wolf, G.; Ruebsamen-Waigmann, H.;  
 Oncogene 8, 3433-3440, 1993

A:Title: Structure, expression and chromosomal mapping of TKT from man and mouse: a n  
 A:Reference number: S42621; MUID:94067796  
 A:Accession: S42621  
 A:Molecule type: mRNA  
 A:Residues: 1-855 <KAR>  
 A:Cross-references: EMBL:X74764; NID:g433337; PID:g433338  
 C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; discolidin I amino-  
 C:Keywords: phosphotransferase; tyrosine-specific protein kinase  
 F:29-185/Domain: discolidin I amino-terminal homology <DNI>  
 F:561-855/Domain: protein kinase homology <KIN>  
 F:569-577/Region: protein kinase ATP-binding motif

Query Match 49.3%; Score 2415; DB 2; Length 855;  
 Best Local Similarity 52.1%; Pred. No. 1.2e-126;  
 Matches 482; Conservative 118; Mismatches 227; Indels 98; Gaps 15;

QY 3 PEALSSLLLLLVASGDADMKGHFDPKACRYALGMDRTIPDSISASSWSDDSTAAHRS 62  
 Db 5 PRMLLVLLLLPILS---SAKAQVNPACRYPLGMSGGOIPDEDITASSOWSESTAAYK 61  
 QY 63 RLESSDGDGAWCPAGSPVPEKEEYLOVDLQRLHLVALVGTQGRHAGGLGKEFSRYSR 121  
 Db 62 RLESSDGDGAWCPVPEPDLLKEFLQIDLHTLFTLVGTQGRHAGGLGKEFSRYSR 121  
 QY 122 YSRDGRWGWKDRWGQEVISGNEDEPGVVLKDLGPPMVARLYRFPYPRADRVMSVCLR 181  
 Db 122 YSRDGRWGWKDRWGQEVISGNEDEPGVVLKDLGPPMVARLYRFPYPRADRVMSVCLR 181  
 QY 182 LYGCLWRDGLLSYAPVGTQMTYSEAYLYNDSTYDGHVTCGLQYGLGLQADGVGLD 239  
 Db 182 LYGCLWRDGLLSYAPVGTQMTYSEAYLYNDSTYDGHVTCGLQYGLGLQADGVGLD 240  
 QY 240 DFKSKOELRWPCYDYGWNSHFSFGYVEMEFEDRLRAFQAMQVHCNNHHTILGARLPG 299  
 Db 241 DFKSKOELRWPCYDYGWNSHFSFGYVEMEFEDRLRAFQAMQVHCNNHHTILGARLPG 299  
 QY 300 GVECFRRGPNAMWEGEPHRLHGNLGDPRARAVSVPLGGVRARFLQCRFLFAGPWLFF 359  
 Db 301 GVECFRRGPNAMWEGEPHRLHGNLGDPRARAVSVPLGGVRARFLQCRFLFAGPWLFF 359  
 QY 360 SEISFISDVVNSSPALGTFEPAPWPPGPPPTNFSSLELEPRCQOQVAKPEGSPTAIL 418  
 Db 360 SEISFISDVVNSSPALGTFEPAPWPPGPPPTNFSSLELEPRCQOQVAKPEGSPTAIL 400  
 QY 419 IGLVAIILLLLLIIALMLRHLWRLLSKAERRVLEELTVHLSVPGDTILNNR--P 476  
 Db 401 IGLVAIILLLLLIIALMLRHLWRLLSKAERRVLEELTVHLSVPGDTILNNR--P 460  
 QY 477 GPREP-----PPYQEPFRGNPNPHSPAPVNGSALLSNPAYRLLATYARP 523  
 Db 461 SPSEQGSNSTYDRIFPLRPDYQEP-----SRLKLPF----- 495  
 QY 524 PRGPGPTTAWAKPTNTQAYSGDYMEPEKPGAPILLPPPPQNSVPHYADIVTLQGTG 583  
 Db 495 -----APGEESGCGVWKPQVSGP-----EGVPHYADIVTLQGTG 535  
 QY 584 NYAVPALPGAVGDGPPRV-DPFRSLRFLKELGEGQFGEVHLCVDSQDLVSLDFPL 642  
 Db 536 NYAVPALPGAVGDGPPRV-DPFRSLRFLKELGEGQFGEVHLCVDSQDLVSLDFPL 595  
 QY 643 NVKRGHPLLVAVKILRPDATKNARNDFLKEVKIMSLRDKPNIIRLLGVCVQDDPLCMITD 702  
 Db 596 DVSANQVPLVAVKMLRADANKNARNDFLKEVKIMSLRDKPNIIRLLGVCVQDDPLCMITD 655  
 QY 703 YMEGDLNQLSAHQLEDAKAGAPGDGQAAGQPTISYPMILHLVAAQIASGMRYLATLNF 762  
 Db 656 YMEGDLNQLSAHQLEDAKAGAPGDGQAAGQPTISYPMILHLVAAQIASGMRYLATLNF 706  
 QY 763 VHRDLATRNCLVGENFTIKIADFGMSRLNYAGDYRVQGRAVLPPIRMWAWECILMKFTT 822  
 Db 707 VHRDLATRNCLVGENFTIKIADFGMSRLNYAGDYRVQGRAVLPPIRMWAWECILMKFTT 766

---

A:Reference number: A44547

A:Accession: A44547

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-248 <REL>

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homology F:1-246/Domain: protein kinase homology (fragment) <KIN>

Query Match 19.4%; Score 952; DB 2; Length 248;

Best Local Similarity 72.3%; Pred. No. 2.9e-46;

Matches 172; Conservative 29; Mismatches 31; Indels 6; Gaps 1;

QY 665 ARNDFELKEVKINSLKDPNIRLLGVGVDDPLCMITDYMNGDLNQFLSAHQLEDKAAE 724

Db 5 SNDFELKELKILSLSDPHIRLLGACLEDPLCMITEYMNGDLNQFLSSHLDEGEN 64

QY 725 GAPGCGAAGQPTISYPMLLHVAQAQASGMRYLATLNFVHRDLATRNCLVGNFTIKTAD 784

Db 65 G-----AHCLPAISYSSLLHVALQICSGMKYLSSLFVHRDLARNCLVGNFTIKTAD 118

QY 785 FGMRNLVAGDYRVGQRAVLPFRMMAWECILMGKFTTASDYWAFGVTILWEVLMCLRAQP 844

Db 119 FGMRNLVAGDYRIQGRAVLPFRMMAWECILMGKFTTASDYWSEGVTLWEILMLCKEQP 178

QY 845 FGLTDEQVIEAGFEFRDQGRVILSRPPACPGLYELMLRCWSEQRPPFSQLH 902

Db 179 YGELSDVEDVIEAGFEFRDQGRVILSRPPACPGLYELMLRCWSEQRPPFSQLH 902

RESULT 8

I59442

receptor tyrosine kinase - rat (fragment)

C:Species: Rattus sp. (rat)

C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 31-Oct-1997

C:Accession: I59442

R:Sakuma, S.; Saya, H.; Ijichi, A.; Tofilon, P.J.

Radiat. Res. 143, 1-7, 1995

A:Title: Radiation induction of the receptor tyrosine kinase gene ptk-3 in normal rat as

A:Reference number: I59442; MUID:95320273

A:Accession: I59442

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-183 <RES>

A:Cross-references: GB:S7556; NID:9957334; PID:9957335

C:Genetics:

C:Gene: Ptk-35

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homology

Query Match

Best Local Similarity 18.9%; Score 924.5; DB 2; Length 183;

Matches 179; Conservative 0; Mismatches 4; Indels 37; Gaps 1;

QY 404 GOQVAKPEGSPAILIGCLVAIIILLIITLMLRWLRLLSKAERRVLEELTVHL 463

Db 1 GOQVAKAEGSPAILIGCLVAIIILLIITLMLRWLRLLSKAERRVLEELTVHL 60

QY 464 SVPGDTILINRRPGPREPPYQEPGRNPPHSAFCVPNGSALLSNPAYRLLATYARP 523

Db 61 SVPGDTILINRRPGPREPPYQEPGRNPPHSAFCVPNGSALLSNPAYRLLATYARP 523

QY 524 PRGPGPTPAWAKPTNTQAYSDYMEPEKPGAPLLPPPPQNSVPHYAEADIVTLQGVGG 583

Db 104 -----SDGYMEPEKPGAPLLPPPPQNSVPHYAEADIVTLQGVGG 143

QY 584 NTYAVPALPPGAVGDPGRVDFPRSRRLRFKEKLGEGQFGE 623

Db 144 NTYAVPALPPGAVGDPGRVDFPRSRRLRFKEKLGEGQFGE 183

RESULT 9

S44098

brain-derived neurotrophic factor receptor precursor - chicken

N:Alternate names: receptor tyrosine kinase trkB

C:Species: Gallus gallus (chicken)

C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 21-Nov-1997

C:Accession: S59339; S42175; S44098

R:Vinh, N.O.; Erdmann, K.S.; Heumann, R.

Gene 149, 383-384, 1994

A:Title: Cloning and sequence analysis of a cDNA encoding a novel truncated form of t

A:Reference number: S59338; MUID:95047511

A:Accession: S59339

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: mRNA

A:Residues: 1-818 <VI2>

A:Cross-references: EMBL:X77251; NID:9563881; PID:g472934

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1994

R:Dechant, G.; Biffo, S.; Okazawa, H.; Kolbeck, R.; Pottgiesser, J.; Barde, Y.A.

Development 119, 545-558, 1993

A:Title: Expression and binding characteristics of the BDNF receptor chick trkB.

A:Reference number: S42175; MUID:94116452

A:Accession: S42175

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-818 <DEC>

A:Cross-references: EMBL:X74109; NID:g407798; PID:g407799

C:Genetics:

A:Gene: trkB

C:Function:

A:Description: regulation of nervous system development; receptor for brain-derived n

C:Superfamily: nerve growth factor receptor, high affinity; leucine-rich alpha-2-glyc

C:Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; growth fact

inase

F:1-31/Domain: signal sequence #status predicted <SIG>

F:32-818/Product: brain-derived neurotrophic factor receptor #status predicted <MAT>

F:32-432/Domain: extracellular #status predicted <EXT>

F:66-90/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>

F:91-114/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>

F:115-137/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>

F:138-149/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>

F:433-449/Domain: leucine-rich alpha-2-glycoprotein repeat homology #status atypical

F:450-818/Domain: transmembrane #status predicted <TMN>

F:832-810/Domain: cytosolic #status predicted <CYT>

F:832-810/Domain: protein kinase homology <KIN>

F:832-810/Region: protein kinase ATP-binding motif

F:832-810/Region: protein kinase ATP-binding motif

F:832-810/Region: protein kinase ATP-binding motif

F:832-810/Region: protein kinase ATP-binding motif

F:832-810/Region: protein kinase ATP-binding motif

F:832-810/Region: protein kinase ATP-binding motif

F:832-810/Region: protein kinase ATP-binding motif

F:832-810/Region: protein kinase ATP-binding motif

F:832-810/Region: protein kinase ATP-binding motif

F:832-810/Region: protein kinase ATP-binding motif

F:832-810/Region: protein kinase ATP-binding motif

F:832-810/Region: protein kinase ATP-binding motif

F:832-810/Region: protein kinase ATP-binding motif

F:832-810/Region: protein kinase ATP-binding motif

F:832-810/Region: protein kinase ATP-binding motif

F:832-810/Region: protein kinase ATP-binding motif

F:832-810/Region: protein kinase ATP-binding motif

F:832-810/Region: protein kinase ATP-binding motif

F:832-810/Region: protein kinase ATP-binding motif

F:832-810/Region: protein kinase ATP-binding motif

F:832-810/Region: protein kinase ATP-binding motif

F:832-810/Region: protein kinase ATP-binding motif

F:832-810/Region: protein kinase ATP-binding motif

F:832-810/Region: protein kinase ATP-binding motif

F:832-810/Region: protein kinase ATP-binding motif

F:832-810/Region: protein kinase ATP-binding motif

```
QY 701 TDYMGDLNOLSAHQLEDKAAEGAPDGOAAGQ---PTISYPMLLHVAQAISGRMYL 757
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 613 FEYKMGDLNKLRAH-----GPDVLMWAGNRPALITQOMLHIAQIAAGNVYL 663

QY 758 ATLVNHRDLATRNCLGENTIKIADFGMRNLVAGDYVRQGRVLPPIKMAWECILM 817
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 664 ASQHEVHRDLATRNCLGENTLVKIGDGMGRDYVSTDYVRGGHTMLPIRMMPPESIMY 723

QY 818 GFTTASDVWAFGVTLWEVLMCLRAQFGOLTDEQVNIENAGEFFRDQGRVYLSRPPACP 877
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 724 RKFTTESDVMGLVNLWEIFYTGK--OPMYQJLNNEVIECI-----TQGR--VLQRPRTCP 775

QY 878 QGLYELMLRCWSRESEQRPPFSQLHREL 905
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 776 KEVYDMLGCWOREPHMLNIKEIHSLL 803

RESULT 10
S06943
brain-derived neurotrophic factor receptor precursor - mouse
N:Alternate names: receptor tyrosine kinase trkB
N:Contains: protein-tyrosine kinase (EC 2.7.1.112)
C:Species: Mus musculus (house mouse)
C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 21-Nov-1997
C:Accession: S06943
EMBO J. 8, 3701-3709, 1989
R:Klein, R.; Parada, L.F.; Coullier, F.; Barbacid, M.
A:Title: trkB, a novel tyrosine protein kinase receptor expressed during mouse neural de
A:Reference number: S06943; MUID:90059970
A:Accession: S06943
A:Molecule type: mRNA
A:Residues: 1-821 <KLE>
A:Cross-references: EMBL:X17647; NID:g55505; PID:g55506
A:Gene: trkB
C:Function:
A:Description: regulation of nervous system development; receptor for brain-derived neur
C:Superfamily: nerve growth factor receptor, high affinity; leucine-rich alpha-2-glycop
C:Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; growth factor
inase
F:1-31/Domain: signal sequence #status predicted <SIG>
F:32-821/Product: brain-derived neurotrophic factor receptor #status predicted <MAT>
F:32-434/Domain: extracellular #status predicted <EXT>
F:67-91/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F:92-115/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F:116-138/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F:139-150/Domain: leucine-rich alpha-2-glycoprotein repeat homology #status atypical <LR
F:435-451/Domain: transmembrane #status predicted <TMN>
F:452-821/Domain: cytosolic #status predicted <CYT>
F:543-551/Region: protein kinase ATP-binding motif
F:67,95,121,178,205,241,254,280,325,338,350,411/Binding site: carbohydrate (Asn) (coval
F:571/Active site: Lys #status predicted
F:706/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicte
F:816/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 13.5%; Score 659.5; DB 1; Length 821;
Best Local Similarity 30.4%; Pred. No. 1.6e-29;
Matches 201; Conservative 72; Mismatches 188; Indels 201; Gaps 22;

QY 373 SPALGCTFPAPWPPGPPPTNFSSLELEPRGQOPVAKPEGSPTAILGLVAILLLLL 432
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 217 SCVGGDPLTLVMDVNLVSKHNNETSHTQSGSLRITNSSDSDSGKQISCV----- 268

QY 433 IIALMLRLHWRLLSKAERVLLEELTVHLSV---PGDTILNNRPPGPREP----- 483
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 268 -----AENLVGEDQDSVNLTVHFAPTITEL-----ESPTSDHHC 302

QY 483 -PYQEPGRGNPPHSAPCPVNGS-----ALLLSNPA-----Y 513
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 303 IFF---TVRGNPKALQFWNGAILNESKYICTKIHVNTHTYHGCLQDNDPTHMNGDY 359
```

```
QY 514 RLL-----ATYARPPRGPGPPPTANAKPTINTQAYSGDYMEP--E 551
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 360 TLMAKNECKDEROISAHFPMGRPGVDYETNPYPEVLYEDWTTPTDI-----GDTTNKSN 415

QY 552 KPGAPLLPPPPQNSPHYAEADIVTLOGV----- 581
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 416 IPSTDVADQSNREHLSVYAVVIVASVVGFCLLVMLLLLLKLARHSKFGMKGPASVLSNDD 475

QY 581 -----TGGNTYAVPALPGGAVGDGP---PRVDFP-----RS 608
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 476 SASPLHISNGSNTSSSEGGPDVLIIGTKIPVIENQYFOITNSQLKPDFTVOHIKRH 535

QY 609 RLRFKEKLGEGOFGEVHLCE---VDSQDLVSLDFPLNVKRGHPLLVAVKILRDPATKNA 665
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 536 NIVLRELGEAGFKVFLAECYNLCPEQD-----KILVAVKTLK-DASDNA 580

QY 666 RNDFLKEVKIMSLKDPNLIIRLLGVCVODPLCLMTIDYMENGLAQFSAHOLED-KAAE 724
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 581 RKDFHREAELLTLQHEHIVKFGVCGEGLIMFVYMKHGLDNKFLRAHGPDAVLNME 640

QY 725 GAPDGOAAQGGPT-ISYPMLLHVAQAISGRMYLATLVNHRDLATRNCLGENTIKIA 783
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 641 GNP-----PTELTQSMQLHIAQIAAGVYLSAQHFVHRDLATRNCLGENTLVKIG 692

QY 784 DFGMSRLNLAGDYIRVQGRVLPPIRMAWECILMGKFTTASDVWAFGVTLWEVLMCLRAQ 843
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 693 DFGMSRDVYSTDYVRGGHTMLPIRMMPPESIMYRKFTTESDVMGLVNLWEIFYTGK-Q 751

QY 844 PFGQLTDEQVNIENAGEFFRDQGRVYLSRPPACPQGLYELMLRCWSRESEQRPPFSQLHR 903
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 752 PWYQLSNNEVIECI-----TQGR--VLQRPRTCPQVYELMLGCWOREPHTRKNISHT 804

QY 904 FL 905
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 805 LL 806

RESULT 11
A39667
brain-derived neurotrophic factor receptor precursor - rat
N:Alternate names: receptor tyrosine kinase trkB,FL
N:Contains: protein-tyrosine kinase (EC 2.7.1.112)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 27-Nov-1991 #sequence_revision 27-Nov-1991 #text_change 21-Nov-1997
C:Accession: A39667
R:Middlemas, D.S.; Lindberg, R.A.; Hunter, T.
Mol. Cell. Biol. 11, 143-153, 1991
A:Title: trkB, a neural receptor protein-tyrosine kinase: evidence for a full-length
A:Reference number: A39667; MUID:91094826
A:Accession: A39667
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-821 <MID>
A:Cross-references: GB:M55291; NID:g207473; PID:g207474
C:Function:
A:Description: regulation of nervous system development; receptor for brain-derived n
C:Superfamily: nerve growth factor receptor, high affinity; leucine-rich alpha-2-glyc
C:Keywords: alternative splicing; ATP; autophosphorylation; glycoprotein; growth fact
inase
F:1-31/Domain: signal sequence #status predicted <SIG>
F:32-821/Product: brain-derived neurotrophic factor receptor #status predicted <MAT>
F:32-434/Domain: extracellular #status predicted <EXT>
F:67-91/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F:92-115/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F:116-138/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F:139-150/Domain: leucine-rich alpha-2-glycoprotein repeat homology #status atypical
F:435-451/Domain: transmembrane #status predicted <TMN>
F:452-821/Domain: cytosolic #status predicted <CYT>
F:543-551/Region: protein kinase ATP-binding motif
F:67,95,121,178,205,241,254,280,325,338,350,411/Binding site: carbohydrate (Asn) (cov
F:571/Active site: Lys #status predicted
F:705/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predi
```

F:816/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 13.5%; Score 659.5; DB 1; Length 821;  
Best Local Similarity 30.4%; Pred. No. 1.6e-29;  
Matches 201; Conservative 72; Mismatches 188; Indels 201; Gaps 22;

QY 373 SPALGTEPPAPWPGPPPTNFSSLELEPRGQQQVAKPEGSPTAILGICLVAILLLLL 432  
DB 217 SCVSGDPLTLYWDVGNLVSKHMMNETSLTQSLRITNISDDSGKQISCV----- 268  
QY 433 IIALMLRHLWRLLSKARRVLEELTVHLSV---PGDTILNNRPGPREPP----- 483  
DB 268 -----AENLVGEDQDSVNLTVHFAPTITL-----ESPSTDHHC 302  
QY 483 -PYQPRPRGNPHSAPCPVNGS-----ALLSNPA-----Y 513  
DB 303 IPF---TVRGNPKPALQFYNGAILNESKYICTKHVTHHTYHGCCLQDNPHTHNGDY 359  
QY 514 RLL-----ATYARPPRGPPPTPAWAKPTNTQAYSGDYNEP--E 551  
DB 360 TLMKANEYKDKQISAHFMGRPGVDYETNPPEVLYEDWTTPTDI----GDTNKSNE 415  
QY 552 KPGAPLLPPPPONSYPHVAEDIVTQGV-----KILVAVTKL-DASDNA 580  
DB 416 IPSTDVADQTNREHLSVYAVVVIASVGVFCILLMLLLKLARHSGKMGKSPASVISNDDD 475  
QY 581 -----TGNTRYAVPALPGVAGDGP---PRVDFF-----RS 608  
DB 476 SASPLHHISNGSNTSPSEGGPDVAVIIGMTKIPVIENQYFGIINSQKLPDTFVQHKR 535  
QY 609 RLRFKEKLGEGGEGVHLC-----VDSPDLVSLDFPLNVRKHPHLLVAVKILRPDATKNA 665  
DB 536 NIVLKRLEGGAGFVLAECYNLCPEQD-----KILVAVTKL-DASDNA 580  
QY 666 RNDFLKEVKIMSLKDPNIIRLLGVCVQDDPLCMITDYMENGLNQFLSAHQLED-KAAE 724  
DB 581 RKDFHREALLTLNLOHEHIVKFGVCGVGDPLIMVFEYMKHGDNLKFLRAHGPDAVLMAE 640  
QY 725 GAPGQGAQAQGT-ISYPMLLHVAQAASGRYLAATLNFVHRDLATRNCLVGENFTIKIA 783  
DB 641 GNP-----PTELTQSMHLIAQQAAGVMYLASQHFVHRDLATRNCLVGENLLVKIG 692  
QY 784 DFGMSRLNLYAGDYRYVQGRAVLPFRMAWECILMGKFTTASDVAFGVTWLWVLMCLRAQ 843  
DB 693 DFGMSRDVYTDYRVGGHTMLPFRMPPESIMYRKFTTESDVWSGLVGLWEIYFTYK-Q 751  
QY 844 PFGQLTDEQVIENAGEFFRDGQGVYLSRPPACPGQGLYELMLRCWSRESQRPPFSOLHR 903  
DB 752 PWYQLSNNEVIECI-----TQGR--VLQRPRTCPQEVYELMLGCGWQREPHTRKNKIHT 804  
QY 904 FL 905  
DB 805 LL 806

RESULT 12  
A56853

brain-derived neurotrophic factor receptor precursor - human  
N:Alternate names: receptor tyrosine kinase trkb  
N:Contains: protein-tyrosine kinase (EC 2.7.1.112)  
C:Species: Homo sapiens (man)  
C:Date: 11-Aug-1995 #sequence\_revision 11-Aug-1995 #text\_change 21-Nov-1997  
C:Accession: A56853; 156557  
R:Nakagawara, A.; Liu, X.G.; Ikegaki, N.; White, P.S.; Yamashiro, D.J.; Nycum, L.M.; Bie  
Genomics 25, 538-546, 1995  
A:Title: Cloning and chromosomal localization of the human TRK-B tyrosine kinase recepto  
A:Reference number: A56853; MUID:95309922  
A:Accession: A56853  
A:Molecule type: mRNA  
A:Residues: 1-822 <NAK>  
A:Cross-references: GB:U12140; NID:g525313; PID:g530791  
R:Shelton, D.L.; Sutherland, J.; Gripp, J.; Camerato, T.; Armanini, M.P.; Phillips, H.S.

J. Neurosci. 15, 477-491, 1995  
A:Title: Human trks: molecular cloning, tissue distribution, and expression of extrac  
A:Reference number: 156557; MUID:95123473  
A:Accession: 156557  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: mRNA  
A:Residues: 1-822 <SHE>  
A:Cross-references: GB:S76473; NID:g913717; PID:g913718  
C:Genetics:  
A:Gene: GDB:NTRK2; trkb  
A:Cross-references: GDB:127898; OMIM:600456  
A:Map position: 9q22.1-9q22.1  
C:Function:  
A:Description: regulation of nervous system development; receptor for brain-derived n  
C:Superfamily: nerve growth factor receptor, high affinity; leucine-rich alpha-2-glyc  
C:Keywords: alternative splicing; AIP; autophosphorylation; glycoprotein; growth fact  
inase

F:1-31/Domain: signal sequence #status predicted <SIG>  
F:32-822/Product: brain-derived neurotrophic factor receptor #status predicted <MAT>  
F:32-435/Domain: extracellular #status predicted <EXT>  
F:97-91/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>  
F:92-115/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>  
F:116-138/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>  
F:139-150/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>  
F:436-452/Domain: transmembrane #status predicted <TMN>  
F:453-822/Domain: cytosolic #status predicted <CYT>  
F:536-814/Domain: protein kinase homology <KIN>  
F:544-552/Region: protein kinase AIP-binding motif  
F:57,95,121,178,204,234,280,325,338,350,412/Binding site: carboxydrate (Asn) (cov  
F:572/Active site: Lys #status predicted  
F:706/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predi  
F:817/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 13.4%; Score 658; DB 1; Length 822;

Best Local Similarity 47.4%; Pred. No. 2e-29;  
Matches 144; Conservative 45; Mismatches 79; Indels 36; Gaps 9;

QY 607 RSLRFEKLGEGGEGVHLC-----VDSPDLVSLDFPLNVRKHPHLLVAVKILRPDATK 663  
DB 535 RHNVILKRELGGAGFVLAECYNLCPEQD-----KILVAVTKL-DASD 579  
QY 664 NARNDLKEVKIMSLKDPNIIRLLGVCVQDDPLCMITDYMENGLNQFLSAHQLED-KA 722  
DB 580 NARKDPHREALLTLNLOHEHIVKFGVCGVGDPLIMVFEYMKHGDNLKFLRAHGPDAVL 639  
QY 723 AEGAPGQGAQAQGT-ISYPMLLHVAQAASGRYLAATLNFVHRDLATRNCLVGENFTIK 781  
DB 640 AEGNP-----PTELTQSMHLIAQQAAGVMYLASQHFVHRDLATRNCLVGENLLVK 691  
QY 782 IADFGMSRLNLYAGDYRYVQGRAVLPFRMAWECILMGKFTTASDVAFGVTWLWVLMCLR 841  
DB 692 IGDGMSRDVYTDYRVGGHTMLPFRMPPESIMYRKFTTESDVWSGLVGLWEIYFTYK 751  
QY 842 AQPFGQLTDEQVIENAGEFFRDGQGVYLSRPPACPGQGLYELMLRCWSRESQRPPFSOL 901  
DB 752 -QPWYQLSNNEVIECI-----TQGR--VLQRPRTCPQEVYELMLGCGWQREPHTRKNKI 803  
QY 902 HREL 905  
DB 804 HTLL 807

RESULT 13

TVRRTB  
nerve growth factor receptor precursor, high affinity - black rat  
N:Alternate names: receptor tyrosine kinase trka  
N:Contains: protein-tyrosine kinase (EC 2.7.1.112)  
C:Species: Rattus rattus (black rat, roof rat)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 21-Nov-1997  
C:Accession: A41981  
R:Weakin, S.O.; Suter, U.; Drinkwater, C.C.; Weicher, A.A.; Shooter, E.M.  
Proc. Natl. Acad. Sci. U.S.A. 89, 2374-2378, 1992

A:Title: The rat trk protooncogene product exhibits properties characteristic of the slc  
A:Reference number: A41981; MUID:92196121  
A:Accession: A41981  
A:Molecule type: mRNA  
A:Residues: 1-799 <MEA>  
A:Cross-references: GB:M85214; NID:g207481; PID:g207482  
A:Note: sequence extracted from NCBI backbone (NCBI:88433, NCBI:88434)  
C:Comment: The proto-oncogene trkA is activated by gene fusion. The amino end of several  
C:Function:  
A:Description: regulation of nervous system development; receptor for nerve growth factor  
C:Superfamily: nerve growth factor receptor, high affinity; leucine-rich alpha-2-glycoprotein  
C:Keywords: ATP; autophosphorylation; glycoprotein; growth factor receptor; phosphoprotein  
C:Keywords: ATP; autophosphorylation; glycoprotein; growth factor receptor; phosphoprotein  
F:1-32/Domain: signal sequence #status predicted <SIG>  
F:33-799/Product: nerve growth factor receptor, high-affinity #status predicted <MAT>  
F:33-424/Domain: extracellular #status predicted <EXT>  
F:67-91/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>  
F:92-115/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>  
F:116-138/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>  
F:139-150/Domain: leucine-rich alpha-2-glycoprotein repeat homology #status atypical <LR>  
F:425-441/Domain: transmembrane #status predicted <TMN>  
F:442-799/Domain: cytosolic #status predicted <CYT>  
F:511-791/Domain: protein kinase homology <KIN>  
F:519-527/Region: protein kinase ATP-binding motif  
F:67-121,190,204,255,264,320,325,341,361,404/Binding site: carbohydrate (Asn) (covalent)  
F:547/Active site: Lys #status predicted  
F:683/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted  
F:794/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 13.3%; Score 653.5; DB 1; Length 799;  
Best Local Similarity 29.8%; Pred. No. 3,4e-29;  
Matches 190; Conservative 78; Mismatches 155; Indels 215; Gaps 21;

QY 335 SVPLGGRVARELOC-----RFLFAGPWLLESEISFI----- 366  
DB 289 SVHLGKAEQHHWCIPFSDGQAPSLRWFFNGS--VLNLSFTFTFLSALTNETMRH 346  
QY 366 -----SDVYNNSSPALGGTFP-----PAPWMPGPPPTNFSLELEPR 403  
DB 347 GCLRLNQTHVNGNYTLAANPYGQAASTMAAFMDNPFENFEDPIPVSEFVDNST 406  
QY 404 GOQPVAKPEGSPATLIGCLVAIIILLIILMLRLHWRLLSKAERR-----V 454  
DB 407 SRDPVEKKDETFFGSVAVGLVAALFL-SALL-----VLNKGQSKFGINRAV 458  
QY 455 LEE-----LTVHLSVPGTILNRPGRPPPYQEPGRNPPHSPAPCPVNGSAL--L 507  
DB 459 LAPEDGLAMSLHFMTLGGSSL-----SPTGKSGSLGQHI 493  
QY 508 LSNPAYRLLATYARPPRGPGPTPAWAKPTNTQAYSGDYMEPEKPGAPLLPPPPQNSVP 567  
DB 494 MENPOY-----FS-----DTCVH 506  
QY 568 HYAEADIVTLOGVTGGNTYAVPALPPGAVGDGPRVDFPRSLRKEKLGQGEVHLC 627  
DB 507 HTRQDII-----LKWELGSGAFGKFLA 530  
QY 628 E---VDSQDLVDFPLNVRKGHPLLVAVKILRPDATKNARNDLFEVKIMSLKDPNI 684  
DB 531 ECYNLLNDD-----KMLVAVKALK--ETSENARQDFHREAEILLTQLQHI 575  
QY 685 IRLGLVCVQDDPLCMITDYMENGLDNQFSLAHQLEDKAAEGAPDGOAAQGTYSYPMLL 744  
DB 576 VRFEGVCTEGGLLAVFVYMRHGLNRLRSHGPDAKLAGGE--DVAPGP-LGLGQLL 631  
QY 745 HYAAQTASGRVYLATLNFVHRDLATRNCLVGENFTIKTADFGMGRNLYAGDYRVQGRV 804  
DB 632 AVASQVAGMVLASLHVFHRLATRNCLVGLVGVKIDFGCMRDISTDYRVYRGRTM 691  
QY 805 LPIRMWAEICILMGKFTTASDVAFVTLWEVLMLCRAQPGQLTDEQVIENAGEFFRDQ 864  
DB 692 LPIRMWPPESILYRKFTESDWSVGVVLEIFTYIGK--QPMVQLSNTAEICI-----TQ 745

QY 865 GRQVYLSPRPACPOGLYELMLRCSRESEQRPPFSQLH 902  
DB 746 GRE--LERPRCPDPDYAIMRGWCQREPOQRSLMKDVH 781

RESULT 14  
A55178  
neurotrophin receptor trkC precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 27-Jan-1995 #sequence\_revision 27-Jan-1995 #text\_change 24-Sep-1998  
C:Accession: A55178  
R:McGregor, L.M.; Baylin, S.B.; Griffin, C.A.; Hawkins, A.L.; Nelkin, B.D.  
Genomics 22, 267-272, 1994  
A:Title: Molecular cloning of the cDNA for human TrkC (NTRK3), chromosomal assignment  
A:Reference number: A55178; MUID:95104834  
A:Accession: A55178  
A:Status: Preliminary  
A:Molecule type: mRNA  
A:Residues: 1-825 <MCG>  
A:Cross-references: GB:U05012; NID:g442389; PID:g442390  
C:Genetics:  
A:Gene: GDB:NTRK3  
A:Cross-references: GDB:127899; OMIM:191316  
A:Map position: 19q24-15q25  
C:Superfamily: nerve growth factor receptor, high affinity; leucine-rich alpha-2-glyc  
C:Keywords: alternative splicing; ATP; growth factor receptor  
F:79-103/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>  
F:104-127/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>  
F:128-150/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>  
F:151-162/Domain: leucine-rich alpha-2-glycoprotein repeat homology #status atypical  
F:536-817/Domain: protein kinase homology <KIN>  
F:544-552/Region: protein kinase ATP-binding motif

Query Match 13.3%; Score 651.5; DB 2; Length 825;  
Best Local Similarity 32.9%; Pred. No. 4.6e-29;  
Matches 174; Conservative 66; Mismatches 132; Indels 157; Gaps 18;

QY 490 RGNPPHSPAPCPVNG-----SALLSNPA-----YRLLIATYAR 522  
DB 326 RGNPPTLHLWNGQPLRESKIIHVEYQEGEISGCLLKNKPTHYNNNTLI---AK 381  
QY 523 PPRGPGPPPAWAKTNTQAYSGDYME--PEKGPALL-----PPPNSSVPHVAAEDIV 575  
DB 382 NPLG-----TANQINGHELKEPPESTONFILFDEVSPPTITVTHKPEEDTF 430  
QY 576 TLQGYTGGNTYAVPAL-----P 592  
DB 431 GVSIAVGLAAAFACVLLVLFVMIKYGRRKMGKGPVAVISGEDSDASPLHHNHGITT 490  
QY 593 PGAVGDGP-----PRVDFP-----RSRLRFKEKLGEGQFG 622  
DB 491 PSLDAGPTVVGTRIPVNIENPOYFROGHNCHKPDYVQHIKRRDIVLKRLEGAGFG 550  
QY 623 EYHLCEVDSQDLVSLDFPLNVRKGHPLLVAVKILRPDATKNARNDLFEVKIMSLKDP 682  
DB 551 KVFLAEC-----YNLSPTK-DKMLVAVKALK-DPTLAARKDQREAEELLNLQHE 598  
QY 683 NIIRLLGVCVQDDPLCMITDYMENGLDNQFSLAHQLEDKAAEGAPG-----DGOAAQGP- 737  
DB 599 HIVKFGVCGDGDPLIMVFEYMKHGDNLKFLAH-----GNAMITLVGQPPQKANG 649  
QY 737 TISYPMLLHVAQAISGRMYLATLNFVHRDLATRNCLVGENFTIKIADFGMGRNLYAGDY 796  
DB 650 ELGLSQMLHIAISOIASGMVYLASQHFVHRDLATRNCLVGNALLVKIGDGMGRSDYSTDY 709  
QY 797 YRVQGRVLPPIRMWAEICILMGKFTTASDVAFVTLWEVLMLCRAQPGQLTDSQVTEN 856  
DB 710 YRVGHTMLPIRMWPPESILYRKFTESDWSVGVVLEIFTYIGK--QPMVQLSNTAEIEC 768  
QY 857 AGEFFRDQGRQVYLSRPPACPOGLYELMLRCSRESEQRPPFSQLHREL 905

Db 769 I-----TQGR--VLERPRVCPKEVDVMLGCMQREPOQRNLKEIKIL 810

RESULT 15

A40026

neurotrophin-3 receptor precursor - pig

N:Alternate names: receptor tyrosine kinase trkc

N:Contains: protein-tyrosine kinase (EC 2.7.1.112)

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 17-Jan-1992 #sequence\_revision 17-Jan-1992 #text\_change 21-Nov-1997

C:Accession: A40026

R:Lamballe, F.; Klein, R.; Barbacid, M.

Cell 66, 967-979, 1991

A:Title: trkc, a new member of the trk family of tyrosine protein kinases, is a receptor

A:Reference number: A40026; MUID:91364178

A:Accession: A40026

A:Molecule type: mRNA

A:Residues: 1-825 <L>AM>

A:Cross-references: GB:M80800; NID:g164698; PID:g164699

A:Experimental source: adult brain

C:Function:

D:Description: regulation of nervous system development; receptor for neurotrophin-3

C:Superfamily: nerve growth factor receptor, high affinity; leucine-rich alpha-2-glycoprotein

C:Keywords: ATP; autophosphorylation; brain; glycoprotein; growth factor receptor; phosph

F:1-31/Domain: signal sequence #status predicted <Sig>

F:32-825/Product: neurotrophin-3 receptor #status predicted <MAT>

F:32-436/Domain: extracellular #status predicted <EXT>

F:79-103/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>

F:104-127/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>

F:128-150/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

F:151-162/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

F:437-453/Domain: transmembrane #status predicted <TMN>

F:454-825/Domain: cytosolic #status predicted <CYT>

F:536-817/Domain: protein kinase homology <KIN>

F:544-552/Region: protein kinase ATP-binding motif

F:58,72,79,133,163,203,218,232,259,267,272,294,375,388/Binding site: carbohydrate (Asn)

F:572/Active site: Lys #status predicted

F:709/Binding site: phosphate (Tyr) (covalent) (by autophosphorylation) #status predicted

F:820/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 13.1%; Score 642.5; DB 1; Length 825;

Best Local Similarity 30.4%; Pred. No. 1.4e-28;

Matches 190; Conservative 82; Mismatches 179; Indels 175; Gaps 26;

QY 309 PAMAW--EGEPMRHLGNLGDPRARAVSVPLGGVAREFLQCRFLFAGPWLIFSIFIS 366

Db 331 PTLHLHNGOPLR-----ESKITHVEYQGEVSE--GC-----LLFNKPTHYN 372

QY 367 D---VYNNSSP-----ALGGTFPPAPWPPGPPPT-NFSSL-ELEPRGQOPVA-KPEGGS 414

Db 373 NGNYTLNROEPLGTANQTINGHFLKEPF---PESTDNFVSEYVSPPTITVTHKPEED 428

QY 415 PTAIGLCGLVAIILLLLIALLMLRHLWRLLSKAERRVLEELTVHLSVPGDILINN 474

Db 429 TFCVSTAVGLAFACVLLVLLFIMINKYGR--SK-----FGMKGPVAVISG 473

QY 475 RPPGPREPPPYQEP-----RPRGNPPHSAPCPVNGSALLSNPAYRLLIATYARPPR 525

Db 474 EEDSASPLHHDQPHHHHTLITCRAGHSVIGMTRIP-----VIENPQY-----517

QY 526 GPGPPTPAKAKTNTQAYSGDYMPEKPGAPLLPPPPQNSVPHYAEADIVTLQGVGTGNT 585

Db 517 -----FROGHNCH-----KPDITYVQHKKRRDIV-----540

QY 586 YAVPALPGGAVGDGPRVDFPRSLRFEKGLGEGFVHLCVDSQDLYSLDFPLNVR 645

Db 540 -----LKRELGEAGFKVFLAECYNLS-----PTKVK 566

QY 646 KGHPLLVAVKILRPDATKNARNDLFKEVKIMSKLDPNIRLLGVCVODDPLCMITDYME 705

Db 567 ----MLVAVKALK-DPTLAARKDFQREALLTNLQHEHIVKFYGVCGDGDPLINVFYMK 621

QY 706 NGDINQFLSAHQLEDKAAEGAPG-----DQAAAGQP--TISYPMLLHVAAQIASGRMYLAT 759

Db 622 HGDNLKFLRAH-----GPDAMILVDGQPRQANGELGSLQMLHIASQICSGWYLAS 672

QY 760 LNFVHRDLATRNCLVGENFTIKIADFGMSRNLYAGDYIRVQGRAVLPIRMAWECILMGK 819

Db 673 QHFVHRDLATRNCLVGANLLVKIGDFGMSRDVYSTDYIRVGGHTMLPIRMMPPESIMYRK 732

QY 820 FTTASDVWAFGVTILWEVLMCLCRAQPFQGLTDEQVIENAGFEFFRDGQROVYLSRPPACQOG 879

Db 733 FTTESDVWSEGVILWEIFTYTK-OPWFQLSNTEVIECI-----TQGR--VLERPRVCPKE 784

QY 880 LYELMLRCWSESEQRPPFSQHLRFL 905

Db 785 YVDVMLGCMQREPOQRNLKEIKIL 810

Search completed: November 4, 1999, 01:52:29

Job time: 7584 sec





GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 4, 1999, 05:43:04 ; Search time 752.95 Seconds  
(without alignments)  
9527.996 Million cell updates/sec

Title: US-08-170-558-3  
Perfect score: 3637  
Sequence: 1 GAATTCGTGAGTCGACGTTG.....AAAAAAAAAAAAAAAAAAAA 3637

Scoring table: IDENTITY\_NUC  
Searched: 2546578 seqs, 986266752 residues

Database :

EST:\*

- 1: em\_est1:\*
- 2: em\_est2:\*
- 3: em\_est3:\*
- 4: em\_est4:\*
- 5: em\_est5:\*
- 6: em\_est6:\*
- 7: em\_est7:\*
- 8: em\_est8:\*
- 9: em\_est9:\*
- 10: em\_est10:\*
- 11: em\_est11:\*
- 12: em\_est12:\*
- 13: em\_est13:\*
- 14: em\_est14:\*
- 15: em\_est15:\*
- 16: em\_est16:\*
- 17: em\_est17:\*
- 18: em\_est18:\*
- 19: em\_est19:\*
- 20: gb\_est1:\*
- 21: gb\_est2:\*
- 22: gb\_est3:\*
- 23: gb\_est4:\*
- 24: gb\_est5:\*
- 25: gb\_est6:\*
- 26: gb\_est7:\*
- 27: gb\_est8:\*
- 28: gb\_est9:\*
- 29: gb\_est10:\*
- 30: gb\_est11:\*
- 31: gb\_est12:\*
- 32: gb\_est13:\*
- 33: gb\_est14:\*
- 34: gb\_est15:\*
- 35: gb\_est16:\*
- 36: gb\_est17:\*
- 37: gb\_est18:\*
- 38: gb\_est19:\*
- 39: gb\_est20:\*
- 40: gb\_est21:\*
- 41: gb\_est22:\*
- 42: gb\_est23:\*
- 43: gb\_est24:\*
- 44: gb\_est25:\*
- 45: gb\_est26:\*
- 46: gb\_est27:\*
- 47: gb\_est28:\*
- 48: gb\_est29:\*
- 49: gb\_est30:\*
- 50: gb\_est31:\*
- 51: gb\_est32:\*
- 52: em\_est20:\*
- 53: em\_est21:\*

- 54: em\_est22:\*
- 55: em\_est23:\*
- 56: em\_est24:\*
- 57: em\_est25:\*
- 58: em\_est26:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	638	17.5	652	49	AL047539	AL047539 DKF2p586C
C 2	572.4	15.7	772	54	HS0008101	AL043251 Homo sapi
C 3	505.4	13.9	507	45	A1342378	A1342378 qt27e10.x
C 4	493.4	13.6	495	54	HS0010070	A1045220 Homo sapi
C 5	475.8	13.1	481	38	AA781458	AA781458 aj26f12.s
C 6	471.6	13.0	519	31	AA325140	AA325140 EST28097
C 7	470.8	12.9	474	43	A1204173	A1204173 qd70e12.x
C 8	457.6	12.6	464	45	A1365945	A1365945 ao91a04.x
C 9	448.4	12.3	532	28	AA116659	AA116659 mq23f05.r
C 10	444.4	12.2	446	44	A1275058	A1275058 ql72d06.x
C 11	440.4	12.1	453	39	AA831559	AA831559 oc84d05.s
C 12	432.2	11.9	542	40	AA944928	AA944928 EST200427
C 13	416.4	11.4	429	33	AA430320	AA430320 zw07f11.r
C 14	416.2	11.4	422	44	A1280882	A1280882 qk47b03.x
C 15	416	11.4	465	27	C04008	C04008 C04008 Huma
C 16	411.4	11.3	524	42	A1145187	A1145187 UI-R-BT0-
C 17	411.2	11.3	416	48	A1588864	A1588864 tq28d12.x
C 18	408.4	11.2	418	20	T09338	T09338 EST07231 in
C 19	406.4	11.2	408	48	A1567223	A1567223 tp51e07.x
C 20	400.8	11.0	404	43	A1214511	A1214511 qm27g07.x
C 21	396.6	10.9	422	34	AA464652	AA464652 zx85c05.r
C 22	393	10.8	393	47	A1494466	A1494466 qz16a09.x
C 23	392.4	10.8	406	50	A1686734	A1686734 tx09d08.x
C 24	391.8	10.8	406	40	AA938929	AA938929 op83c09.s
C 25	389.4	10.7	505	44	A1323681	A1323681 mq23f05.x
C 26	381	10.5	477	46	A1430426	A1430426 mf77g10.y
C 27	380.4	10.5	382	38	AA808281	AA808281 oc41d07.s
C 28	379.8	10.4	383	20	T09337	T09337 EST07230 in
C 29	378.4	10.4	404	50	A1689485	A1689485 tx94g05.x
C 30	370.2	10.2	440	45	A1326485	A1326485 mq23f05.y
C 31	366.4	10.1	368	38	AA746483	AA746483 nw62e12.s
C 32	364.4	10.0	413	34	AA464551	AA464551 zx85c05.s
C 33	363.8	10.0	378	47	A1476072	A1476072 tl97e09.x
C 34	360.6	9.9	571	20	X73728	X73728 CHESTW025 G
C 35	355.8	9.8	433	46	AA956276	AA956276 UI-R-E1-f
C 36	349.4	9.6	383	33	AA385987	AA385987 EST99705
C 37	348.8	9.6	425	51	A1712520	A1712520 UI-R-AE1-
C 38	347.2	9.5	364	36	AA612719	AA612719 ng26c10.s
C 39	346.8	9.5	361	35	AA574033	AA574033 nkl1a11.s
C 40	344.4	9.5	347	50	A1686242	A1686242 tu40h04.x
C 41	343	9.4	424	32	AA338319	AA338319 EST43238
C 42	335.8	9.2	339	49	A1620626	A1620626 tu47g12.x
C 43	334.8	9.2	730	26	W22881	W22881 75B10 Human
C 44	332.4	9.1	422	26	W98395	W98395 mg20b06.r1
C 45	330	9.1	379	20	T29903	T29903 EST99706 Hu

ALIGNMENTS

RESULT 1  
AL047539/c  
LOCUS AL047539 652 bp mRNA  
DEFINITION DKF2p586C0921.sl 586 (synonym: hute1) Homo sapiens CDNA clone  
ACCESSION DKF2p586C0921, mRNA sequence.  
NID AL047539  
VERSION 94728535  
AL047539.1 GI:4728535

KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
AUTHORS Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Ottenwaelder B., Obermaier, B., Mewes, W., Gassenhuber, J. and  
Wiemann, S.  
TITLE EST (Ottenwaelder, et al.)  
JOURNAL Unpublished (1999)  
COMMENT On Jun 5, 1998 this sequence version replaced gi:3189690.

Contact: Ottenwaelder B

MIPS  
Am Klopferspitz 18a D-82152 Martinsried, Germany.

## FEATURES

source

1. .652  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="DKFZp386C0921"  
/clone\_lib="586 (synonym: hutel)"  
/tissue\_type="uterus"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/note="vector: pSport1; Site\_1: NotI; Site\_2: SalI/MLuI"  
BASE COUNT 161 a 165 c 174 g 152 t  
ORIGIN

Query Match 17.5%; Score 638; DB 49; Length 652;  
Best Local Similarity 99.8%; Pred. No. 4.6e-117;  
Matches 649; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 2959 AGAAGCCCTGTGCGCCACCCAGCTGTCTGTGGATGGATCTCTCCACCCCTCTCTA 3018  
DB 649 AGAAGCCCTGTGCGCCACCCAGCTGTCTGTGGATGGATCTCTCCACCCCTCTCTA 590

QY 3019 GCCATCTCTGGGAAGGGTGGGAGAATATAGATAGACATGGACATGGCCCATGG 3078  
DB 589 GCCATCTCTGGGAAGGGTGGGAGAATATAGATAGAC-CTGGACATGGCCCATGG 531

QY 3079 AGCAGCTGGGCCCCACTGAGCAACACTGATCTCTGGAGAGTGGCTGGCCCGCAGCTTC 3138  
DB 530 AGCAGCTGGGCCCCACTGAGCAACACTGATCTCTGGAGAGTGGCTGGCCCGCAGCTTC 471

QY 3139 TCTCTCCCTGTACACACTGGACCCCACTGGCTGAGAACTCTGGGGTGGAGGACACA 3198  
DB 470 TCTCTCCCTGTACACACTGGACCCCACTGGCTGAGAACTCTGGGGTGGAGGACACA 411

QY 3199 AGGAGAGAAATGTTTCTTGTGCTGCTCTGTACTTGTCTCAGCTTGGGCTTCTTC 3258  
DB 410 AGGAGAGAAATGTTTCTTGTGCTGCTCTGTACTTGTCTCAGCTTGGGCTTCTTC 351.

QY 3259 CTCCTCCATCACTGAACACTGGACCTGGGGGTAGCCCGCCCGCCCTCAGTCACCC 3318  
DB 350 CTCCTCCATCACTGAACACTGGACCTGGGGGTAGCCCGCCCGCCCTCAGTCACCC 291

QY 3319 CCACCTCCCACTGCAGCTTGTAGTAGAACTTCTTAAGCCCTATACGTTCTGTGGAG 3378  
DB 290 CCACCTCCCACTGCAGCTTGTAGTAGAACTTCTTAAGCCCTATACGTTCTGTGGAG 231

QY 3379 TAAATATTGGGATTTGGGGGAAAGAGGAGCAACGGCCCATAGCTTGGGGTGGACATC 3438  
DB 230 TAAATATTGGGATTTGGGGGAAAGAGGAGCAACGGCCCATAGCTTGGGGTGGACATC 171

QY 3439 TCTAGTGTAGCTGCCACATGATTTTCTATAATCACTTGGGGTGTGTACATTTTGGGG 3498  
DB 170 TCTAGTGTAGCTGCCACATGATTTTCTATAATCACTTGGGGTGTGTACATTTTGGGG 111

QY 3499 GGAGAGACAGATTTTACACTAATATATAGACCTAGCTTGGAGCAATTTAATCCCT 3558  
DB 110 GGAGAGACAGATTTTACACTAATATATAGACCTAGCTTGGAGCAATTTAATCCCT 51

QY 3559 GCACCTAGGCAGGTAATAATAAGTTGAGTTTCCACAAAAA 3608  
DB 50 GCACCTAGGCAGGTAATAATAAGTTGAGTTTCCACAAAAA 1

## RESULT 2

HSN008101  
ID HSN008101 standard; RNA; EST; 772 BP.

XX AC AL043251;

XX SV AL043251.1

XX NI el1401623

XX DT 12-MAR-1999 (Rel. 59, Created)

XX DT 12-MAR-1999 (Rel. 59, Last updated, Version 1)

DE DE Homo sapiens mRNA; EST DKFZp434K0223\_r1 (from clone  
DE DKFZp434K0223)

XX KW EST; expressed sequence tag.

XX OS Homo sapiens (human)

XX OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria;

XX OC Primates; Catarrhini; Hominidae; Homo.

XX RN [1]

XX RP 1-772

RA Blum H., Bauersachs S., Mewes W., Gassenhuber J., Wiemann S.;

RT ; Submitted (12-MAR-1999) to the EMBL/GenBank/DBJ databases.

RL MIPS, Am Klopferspitz 18a D-82152 Martinsried, GERMANY

XX CC Clone from S. Wiemann, sequenced by LMU within the cDNA

XX CC sequencing consortium of the German Genome Project

XX CC No si sequence available

XX CC This clone is available at the RZPD in Berlin

XX CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

XX CC Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de

XX FH Key

XX FH Location/Qualifiers

FT source

FT 1..772

FT /db\_xref="taxon:9606"

FT /organism="Homo sapiens"

FT /clone="DKFZp434K0223"

FT /clone\_lib="434 (synonym: htes3). Vector pSport1; host

FT DH10B; sites NotI + SalI"

FT /dev\_stage="adult"

FT /tissue\_type="testis"

XX SQ Sequence 772 BP; 118 A; 250 C; 260 G; 143 T; 1 other;

Query Match 15.7%; Score 572.4; DB 54; Length 772;  
Best Local Similarity 99.3%; Pred. No. 5e-104;  
Matches 595; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 38 AGAGATGCTGCCCCCAGCCCTTAGGCCGAGGGATCAGGAGCTATGGACACAGGGCCC 97  
DB 158 AGAGATGCTGCCCCCAGCCCTTAGGCCGAGGGATCAGGAGCTATGGACACAGGGCCC 217

QY 98 TGTCTATCTTTACTGCTGCTCTTGGTGGCAAGTGGAGATGCTGACATGAAGGACATT 157  
DB 218 TGTCTATCTTTACTGCTGCTCTTGGTGGCAAGTGGAGATGCTGACATGAAGGACATT 277

QY 158 TTGATCTCTGCCAAGTGGCCGCTATGCCCTGGGCATGCGAGACCGGACCATCCAGACATG 217  
DB 278 TTGATCTCTGCCAAGTGGCCGCTATGCCCTGGGCATGCGAGACCGGACCATCCAGACATG 337

QY 218 ACATCTCTGCTTCCAGCTCTCTGGTTCAGATTCCACTGCGCCGCCAGACAGGATTTGAGA 277  
DB 218 ACATCTCTGCTTCCAGCTCTCTGGTTCAGATTCCACTGCGCCGCCAGACAGGATTTGAGA 277

```

Db 338 ACATCTCTCTCCAGCTCCTGGTCAGATTCCACTGCCGCCGCCACAGCAGGTTGGAGA 397
QY 278 GCAGTGACGGGATGGGCTGTGTCGCCCGCAGGGTCGGTGTTCCTCCAGGAGGAGGT 337
Db 398 GCAGTGACGGGATGGGCTGTGTCGCCCGCAGGGTCGGTGTTCCTCCAGGAGGAGGT 457
QY 338 ACTTGACGAGTGATACACAGCTGCACCTGGTGGCTCTGGTGGCCACCCAGGAGCGGC 397
Db 458 ACTTGACGAGTGATACACAGCTGCACCTGGTGGCTCTGGTGGCCACCCAGGAGCGGC 517
QY 398 ATGCCGGGGGCTGGGCAAGGAGTTCTCCCGGAGCTACCGGCTGGTGTACTCCCGGATG 457
Db 518 ATGCCGGGGGCTGGGCAAGGAGTTCTCCCGGAGCTACCGGCTGGTGTACTCCCGGATG 577
QY 458 GTCCGGCTGGATGGGCTGGGCAAGGAGTTCTCCCGGAGCTACCGGCTGGTGTACTCCCGGATG 517
Db 578 GTCCGGCTGGATGGGCTGGGCAAGGAGTTCTCCCGGAGCTACCGGCTGGTGTACTCCCGGATG 637
QY 518 ACCCTGAGGAGTGCTGTAAGGAGTTCTCCCGGAGCTACCGGCTGGTGTACTCCCGGATG 577
Db 638 ACCCTGAGGAGTGCTGTAAGGAGTTCTCCCGGAGCTACCGGCTGGTGTACTCCCGGATG 696
QY 578 TCTACCCCGGGCTGACCGGGTCTAGAGCTGTCTGCTGGGTAGAGCTCTATGCTGCTGC 636
Db 697 TCTACCCCGGGCTGACCGGGTCTAGAGCTGTCTGCTGGGTAGAGCTCTATGCTGCTGC 754

RESULT 3
AI342378/c
LOCUS AI342378 507 bp mRNA EST 13-FEB-1999
DEFINITION qt27e10.x1 Soares_pregnant_uterus_NbHPU Homo sapiens cDNA clone
IMAGE:1949226 3, mRNA sequence.
ACCESSION AI342378
NID 94079305
VERSION AI342378.1 GI:4079305
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 507)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On May 18, 1995 this sequence version replaced gi:810993.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1332 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 454.
Location/Qualifiers
1..507
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1949226"
/clone_lib="Soares_pregnant_uterus_NbHPU"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"
/notes="Organ: uterus; Vector: p7T73-Pac; Site:1: Not I;
Site:2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5',
AACTGGAAGAATTCGGCGCGCTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p7T73 vector. Library
went through one round of normalization. Library
constructed by M. Fatima Bonaldo."

```

## FEATURES

source

```

BASE COUNT 136 a 128 c 131 g 112 t
ORIGIN ...

Query Match 13.9%; Score 505.4; DB 45; Length 507;
Best Local Similarity 99.8%; Pred. No. 8.6e-91;
Matches 506; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3087 GGGCCCACTGACCAACTGATTCCTGGAGAGTGCTGGCCCCCAGATTCTCTCTCC 3146
Db 507 GGGCCCACTGACCAACTGATTCCTGGAGAGTGCTGGCCCCCAGATTCTCTCTCC 448
QY 3147 TGTCAACACTGACCCCACTGGCTGAGAACTCTGGGGTGGAGGACAGAGAGAGG 3206
Db 447 TGTCAACACTGACCCCACTGGCTGAGAACTCTGGGGTGGAGGACAGAGAGAGG 388
QY 3207 AAAATGTTCTTGTGCTCTGCTTACTTGTCTAGCTTGGGCTTCTCTCTCTCC 3266
Db 387 AAAATGTTCTTGTGCTCTGCTTACTTGTCTAGCTTGGGCTTCTCTCTCTCC 328
QY 3267 TCACCTGAAACACTGACCTGGGGTAGCCCCCGCCAGCCCTCAGTCACCCCACTCC 3326
Db 327 TCACCTGAAACACTGACCTGGGGTAGCCCCCGCCAGCCCTCAGTCACCCCACTCC 268
QY 3327 CACCTGCAGTCTGTAGTAGAACTTCTTAAGCCTATACGTTTCTGTGGAGTAATAT 3386
Db 267 CACTTGCAGTCTGTAGTAGAACTTCTTAAGCCTATACGTTTCTGTGGAGTAATAT 208
QY 3387 GGAATTGGGGGAAAGAGGAGCAACGGCCCATAGCTTGGGGTTGGACATCTCTAGTGT 3446
Db 207 GGAATTGGGGGAAAGAGGAGCAACGGCCCATAGCTTGGGGTTGGACATCTCTAGTGT 148
QY 3447 AGCTGCCACATTCATTTTCTATATCACTCTGGGGTTGTACATTTTGGGGGAGAGAC 3506
Db 147 AGCTGCCACATTCATTTTCTATATCACTCTGGGGTTGTACATTTTGGGGGAGAGAC 88
QY 3507 ACAGATTTTACACATAATATATGACCTAGCTTGGAGCAATTTTAAATCCCTGCCTAGG 3566
Db 87 ACAGATTTTACACATAATATATGACCTAGCTTGGAGCAATTTTAAATCCCTGCCTAGG 28
QY 3567 CAGGTAATAATAAGGTTGAGTTTCC 3593
Db 27 CAGGTAATAATAAGGTTGAGTTTCC 1

RESULT 4
HSM010070
ID HSM010070 standard; RNA; EST; 495 BP.
AC AL045220;
XX
SV AL045220.1
XX
NI e1403691
XX
DT 12-MAR-1999 (Rel. 59, Created)
DT 12-MAR-1999 (Rel. 59, Last updated, Version 1)
XX
DE Homo sapiens mRNA; EST DKFZp434H0150_r1 (from clone
DE DKFZp434H0150)
XX
KW EST; expressed sequence tag.
XX
OS Homo sapiens (human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria;
OC Primates; Catarrhini; Hominidae; Homo.
XX
RN [1]
RP 1-495
RA Poustka A., Klein M., Mewes W., Gassenhuber J., Wiemann S.;
RT Submitted (12-MAR-1999) to the EMBL/GenBank/DBJ databases.
RL MIPS, Am Klopferspitze 18a D-82152 Martinsried, GERMANY

```

```

XX CC Clone from S. Wiemann, sequenced by DKFZ within the cDNA
CC CC sequencing consortium of the German Genome Project
CC CC No SI sequence available
CC CC This clone is available at the RZPD in Berlin
CC CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
CC CC Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de
XX XX Key Location/Qualifiers
FH FH Key 1. 495
FH FT source
FH FT /db_xref="taxon:9606"
FH FT /organism="Homo sapiens"
FH FT /clone="DKFZp434H0150"
FH FT /clone_lib="434 (synonym: htes3). Vector pSport1; host
FH FT DH10B; sites NotI + SalI"
FH FT /dev_stage="adult"
FH FT /tissue_type="testis"
XX XX Sequence 495 BP; 105 A; 145 C; 142 G; 103 T; 0 other;

Query Match 13.6%; Score 493.4; DB 54; Length 495;
Best Local Similarity 99.8%; Pred. No. 2e-88;
Matches 494; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1716 GGACTATATGAGCTGAGAGCCAGCGCCCGCTTCTGCCCCACCTCCCCAGAACAG 1775
DB 1 GGACTATATGAGCTGAGAGCCAGCGCCCGCTTCTGCCCCACCTCCCCAGAACAG 60
QY 1776 GGTCCCTCATATGCGAGGCTGACATGTTTACCTGCGAGGCTGACCGGGGCAACAC 1835
DB 61 GGTCCCTCATATGCGAGGCTGACATGTTTACCTGCGAGGCTGACCGGGGCAACAC 120
QY 1836 CTATCTGTGCTGACATGCCCCCAGGCGAGTGGGGATGGGCCCCCAGAGTGGATT 1895
DB 121 CTATCTGTGCTGACATGCCCCCAGGCGAGTGGGGATGGGCCCCCAGAGTGGATT 180
QY 1896 CCTCGATCTGACATCCGCTTCAAGAGAGCTTGGCGAGGCGCCAGTTGGGAGGTGCA 1955
DB 181 CCTCGATCTGACATCCGCTTCAAGAGAGCTTGGCGAGGCGCCAGTTGGGAGGTGCA 240
QY 1956 CTTGTGTGAGTGTGACAGCCCTCAAGATCTGTGTCAGTCTTGATTTCCCTTTAATGTGG 2015
DB 241 CTTGTGTGAGTGTGACAGCCCTCAAGATCTGTGTCAGTCTTGATTTCCCTTTAATGTGG 300
QY 2016 TAAGGACACCTTTTCTGCTGAGTGTCTCAAGATCTTACGCGCAGATGCCCAAGATGC 2075
DB 301 TAAGGACACCTTTTCTGCTGAGTGTCTCAAGATCTTACGCGCAGATGCCCAAGATGC 360
QY 2076 CAGGAATGATTTCTTGAAGAGGTGAAGATCATGTGCGAGGCTCAAGGACCCAAACATCAT 2135
DB 361 CAGGAATGATTTCTTGAAGAGGTGAAGATCATGTGCGAGGCTCAAGGACCCAAACATCAT 420
QY 2136 TCGGCTGTGGGGGTGTGTGTGCGAGGACGACCCCTCTGATGATTAAGTACTGACTACATGGA 2195
DB 421 TCGGCTGTGGGGGTGTGTGTGCGAGGACGACCCCTCTGATGATTAAGTACTGACTACATGGA 480
QY 2196 GAACGCGACCTCAA 2210
DB 481 GAACGCGACCTCAA 495

RESULT 5
LOCUS AA781458/c
DEFINITION aJ26f12.s1 Soares_testis_NHT Homo sapiens cDNA clone 1391471 3',
mRNA sequence.
ACCESSION AA781458
NID 92840789
VERSION AA781458.1 GI:2840789
KEYWORDS EST.
SOURCE human.

```

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 481)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1404759.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

Insert Length: 2285 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 480.
Location/Qualifiers
1. 481
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="1391471"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech
Laboratories, Inc., and primed with a Not I - oligo(dT)
primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCCCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 131 a 117 c 120 g 113 t
ORIGIN
Query Match 13.1%; Score 475.8; DB 38; Length 481;
Best Local Similarity 99.6%; Pred. No. 6.3e-85;
Matches 477; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3124 TGGCCCCCAGCTTCTCTCCCTGTACACACTGGACCCCACTGGCTGAGAATCTGGGG 3183
DB 479 TGGCCCCCAGCTTCTCTCCCTGTACACACTGGACCCCACTGGCTGAGAATCTGGGG 420
QY 3184 GTGAGGAGACAAGAGAGAGAAATGTTCTTGTGCTGCTCTGCTCTGCTCTGCTCTC 3243
DB 419 GTGAGGAGACAAGAGAGAGAAATGTTCTTGTGCTGCTCTGCTCTGCTCTGCTCTC 360
QY 3244 AGCTTGGGCTTCTTCTCTCTCCATCACCTGAAACACTGGACCTGGGGGTAGCCGCCGCC 3303
DB 359 AGCTTGGGCTTCTTCTCTCTCCATCACCTGAAACACTGGACCTGGGGGTAGCCGCCGCC 300
QY 3304 AGCCCTCAGTACACCCCACTTCCCACTGCAGTCTTGTAGCTAGAACTTCTCTAAGCCTA 3363
DB 299 AGCCCTCAGTACACCCCACTTCCCACTGCAGTCTTGTAGCTAGAACTTCTCTAAGCCTA 240
QY 3364 TAGCTTCTGTGAGTAATAATTTGGGATTTGGGGGAAAGAGGAGCAAGCCCATAGCC 3423
DB 239 TAGCTTCTGTGAGTAATAATTTGGGATTTGGGGGAAAGAGGAGCAAGCGCTCATAGCC 180
QY 3424 TTGGGGTTGGACATCTCTAGTGTAGCTGGCACATTTGATTTTCTATAATCACTTGGGGTT 3483

```

```

Db 179 TTGGGTTGGACATCTCTAGTGTAGCTGCCACATGATTTTCTATAATCACTTGGGTT 120
QY 3484 TGTACATTTTGGGGGAGAGACACAGATTTTACACTAATATATGGACCTAGCTTGGG 3543
Db 119 TGTACATTTTGGGGGAGAGACACAGATTTTACACTAATATATGGACCTAGCTTGGG 60
QY 3544 CAATTTTAAATCCCTGCACCTAGGAGGTAATTAATGAAGTTGAGTTTCCACAAAAA 3602
Db 59 CAATTTTAAATCCCTGCACCTAGGAGGTAATTAATGAAGTTGAGTTTCCACAAAAA 1

RESULT 6
AA325140 519 bp mRNA EST 20-APR-1997
LOCUS EST28097 Cerebellum II Homo sapiens cDNA 5' end similar to tyrosine
DEFINITION kinase, receptor (GB:L1315), mRNA sequence.
ACCESSION AA325140
NID 91977395
VERSION AA325140.1 GI:1977395
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 519)
Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Georghagen,N.S.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palancas,R.F., McDonald,D.A., Nguyen,D.T., Pelligriano,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M. and Venter,J.C.
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
Nature 377 (5547 Suppl), 3-174 (1995)
96026280
On Nov 29, 1993 this sequence version replaced gi:430338.
Other_ESTs: TH017728
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (http://www.tigr.org/tdb/ngi/ngi.html)
Seq primer: M13 Reverse.
Location/Qualifiers
1. 519
/organism="Homo sapiens"
/db_xref="ATCC (inhost):125647"
/db_xref="taxon:9606"
/clone_lib="Cerebellum II"
/tissue_type="cerebellum"
/dev_stage="adult"
/note="Organ: brain; Vector: pBluescript SK-; Site:1;
EcoRI: Site-2: XhoI"
117 a 145 c 144 g 108 t 5 others
BASE COUNT
ORIGIN

```

```

Query Match 13.0%; Score 471.6; DB 31; Length 519;
Best Local Similarity 97.7%; Pred. No. 4.4e-84;
Matches 507; Conservative 0; Mismatches 9; Indels 3; Gaps 3;

QY 1965 GGTGACAGCCCTCAAGATCTGCTGAGTCTTTGATTTCCCTTAATGTGCGTAAGGACA 2024
Db 1 GGTGACAGCCCTCAAGATCTGCTGAGTCTTTGATTTCCCTTAATGTGCGTAAGGACA 60
QY 2025 CCCTTTGCTGCTGAGTCTCAAGATCTTTAGCGCCAGATGCCACCAAGATGCCAGGAATGA 2084
Db 61 CCCTTTGCTGCTGAGTCTCAAGATCTTTAGCGCCAGATGCCACCAAGATGCCAGGAATGA 120
QY 2085 TTTCTCTGAAAGAGTGAAGATCATGTCGAGGCTCAAGAGCCCAACCAATCATTCGGCTGCT 2144
Db 121 TTTCTCTGAAAGAGTGAAGATCATGTCGAGGCTCAAGAGCCCAACCAATCATTCGGCTGCT 180
QY 2145 GGGCGTGTGTGTCAGGACGACCCCTCTGCTGATGATTACTGACTACATGGAACCGCGCA 2204
Db 181 GGGCGTGTGTGTCAGGACGACCCCTCTGCTGATGATTACTGACTACATGGAACCGCGCA 240
QY 2205 CCTCAACCAAGTTCCTCAGTGCCCGACCCCTCTGCTGATGATTACTGACTACATGGAACCGCGCA 2264
Db 241 CCTCAACCAAGTTCCTCAGTGCCCGACCCCTCTGCTGATGATTACTGACTACATGGAACCGCGCA 300
QY 2265 GGACGGCCAGGCTCGCGAGGGCCCGACCATCATCAGCTACCCCAATGCTGCTCATGT-CGCAG 2323
Db 301 GGACGGCCAGGCTCGCGAGGGCCCGACCATCATCAGCTACCCCAATGCTGCTCATGTGGGCGAG 360
QY 2324 CCAGATCGCTCGCGCATCGCGTATCTGCGGCACACTCAACTTTGTACATCGGGACCTGG 2383
Db 361 CCAGATCGCTCGCGCATCGCGTATCTGCGGCACACTCAACTTTGTACATCGGGACCTGG 420
QY 2384 CCACGGCGAAGTTCCTGAGT-TGGGAAAATTTTCCACATCAAAATCGACAGCTTTGGCATG 2442
Db 421 NCAAGCGGAAGTTCCTGAGT-TGGGAAAATTTTCCACATCAAAATCGACAGCTTTGGCATG 480
QY 2443 AGCGGAAAGTCTTA-TGCTGGGGACTATTACCGTGTGCA 2480
Db 481 AGCGGAAAGTCTTA-TGCTGGGGACTATTACCGTGTGCA 519

RESULT 7
AI204173 474 bp mRNA EST 28-OCT-1998
LOCUS qd70e12.xl Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1734862
DEFINITION 3', mRNA sequence.
ACCESSION AI204173
NID 93756779
VERSION AI204173.1 GI:3756779
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 474)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On Jan 17, 1998 this sequence version replaced gi:2044289.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 675 Std Error: 0.00

```

Seq primer: -40UP from Gibco  
High quality sequence stop: 395.

FEATURES  
source

1. 474  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="2"  
/clone="IMAGE:1734862"  
/clone\_lib="Soares\_testis\_NHT"  
/sex="male"  
/lab\_host="DH10B"  
/note="vector: p773D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories, Inc., and primed with a Not I - oligo(dT) primer [5]."  
TGTACCAACTGTAAGTGGGAGGGCGCCCAATTTTTTTTTTTT 3'.  
Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. Library went through one round of normalization to Cot5, and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 130 a 121 g 107 t

Query Match 12.9%; Score 470.8; DB 43; Length 474;  
Best Local Similarity 99.6%; Pred. No. 6.1e-84;  
Matches 472; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3123 CTGCGCCCCAGCTTCTCTCCCTGTGCACACCTGGACCCACCTGGCTGAGATCTGGG 3182  
DB 474 CTGCGCCCCAGCTTCTCTCCCTGTGCACACCTGGACCCACCTGGCTGAGATCTGGG 415

QY 3183 GGTGAGGAGCAAGAGAGAGAAATGTTTCCTGTGCTGCTGCTGCTGCTGCTGCT 3242  
DB 414 GGTGAGGAGCAAGAGAGAGAAATGTTTCCTGTGCTGCTGCTGCTGCTGCTGCT 355

QY 3243 CAGCTTGGGCTTCTCTCTCCCTGACCTGAAACACTGGACCTGGGGGTAGCCCGGCC 3302  
DB 354 CAGCTTGGGCTTCTCTCTCCCTGACCTGAAACACTGGACCTGGGGGTAGCCCGGCC 295

QY 3303 CAGCCCTCAGTACCCGCCACTTCCACCTGCGAGTCTGTGTAGTAGAACTTCTTAAGCCT 3362  
DB 294 CAGCCCTCAGTACCCGCCACTTCCACCTGCGAGTCTGTGTAGTAGAACTTCTTAAGCCT 235

QY 3363 ATACGTTCTGTGGAGTAATATTTGGGATTGGGGGAAAGAGGAGCAACGCCCATAGC 3422  
DB 234 ATACGTTCTGTGGAGTAATATTTGGGATTGGGGGAAAGAGGAGCAACGCCCATAGC 175

QY 3423 CTTGGGGTTGGACATCTCTAGTGTAGCTGCCACATTTTCTTATATATCACTTGGGGT 3482  
DB 174 CTTGGGGTTGGACATCTCTAGTGTAGCTGCCACATTTTCTTATATATCACTTGGGGT 115

QY 3483 TTGTACATTTTGGGGGAGAGACACAGATTTTACACTAATATATGACCTAGCTTGAG 3542  
DB 114 TTGTACATTTTGGGGGAGAGACACAGATTTTACACTAATATATGACCTAGCTTGAG 55

QY 3543 GCAATTTTAAATCCCTGCACTAGGAGGTAATAATAAGGTTGAGTTTTCACCA 3596  
DB 54 GCAATTTTAAATCCCTGCACTAGGAGGTAATAATAAGGTTGAGTTTTCACCA 1

RESULT 8  
LOCUS AI365945  
DEFINITION 464 bp mRNA EST 07-JAN-1999  
IMAGE:1953198.3' similar to SW-EDDL\_HUMAN Q08345 EPIHELIAL DISCOIDIN DOMAIN RECEPTOR 1 PRECURSOR ; mRNA sequence.  
ACCESSION AI365945  
NID 94125634  
VERSION AI365945.1 GI:4125634  
KEYWORDS EST.

SOURCE  
ORGANISM  
human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE  
AUTHORS  
1 (bases 1 to 464)  
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, N., Lennon, G., Marra, M., Martin, J., Moore, B., Scheilenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.  
WashU-NCI human EST Project  
Unpublished (1997)  
COMMENT  
On Jan 17, 1998 this sequence version replaced gi:2043665.  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.wustl.edu  
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
Possible reversed clone: similarity on wrong strand  
Possible reversed clone: polyT not found  
Seq primer: -40UP from Gibco.  
Location/Qualifiers  
1. 464  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="20"  
/clone="IMAGE:1953198"  
/clone\_lib="Schiller meningioma"  
/sex="female"  
/tissue\_type="meningioma"  
/dev\_stage="72 years"  
/lab\_host="SOLR"  
/note="Organ: brain; Vector: pBluescript SK- (Stratagene); Site\_1: EcoRI; Site\_2: XhoI; Double-stranded cDNA was prepared from human meningioma using primer 5'-GAGAGAGAGAGAGAGAACTAGTGTAGT(18)-3'. An EcoRI adaptor was used on the 5' end of the cDNA as follows: 5'-AATTCGGCAGCAG-3'. The library was size-selected and went through one round of amplification. Average insert size is 1.7 kb, with a range from 0.4-12 kb. Tumor identification by consensus pathology. This library was constructed by Dr. Martin Schiller (Johns Hopkins University)."

BASE COUNT 85 a 134 c 155 g 90 t

FEATURES  
source

Query Match 12.6%; Score 457.6; DB 45; Length 464;  
Best Local Similarity 99.1%; Pred. No. 2.5e-81;  
Matches 460; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2225 CCACACAGCTGGAGGAGCAAGCGAGCGGGGGCCCTGGGGACGGCGAGCTGCGCAGG 2284  
DB 1 CCCACAGCTGGTGGACAGAGCGAGCGGGGGCCCTGGGGACGGCGAGCTTGGCAGG 60

QY 2285 GGCCACCATCAGTACCCCAATGCTGCTGCATGTGGAGCCCCAGATGCGCTCCGCGATGC 2344  
DB 61 GGCCACCATCAGTACCCCAATGCTGCTGCATGTGGAGCCCCAGATGCGCTCCGCGATGC 120

QY 2345 GCTATCTGGCCACACTCAACTTTGTACATCGGACCTGGCCAGCGAACTGCTAGTTG 2404  
DB 121 GCTATCTGGCCACACTCAACTTTGTACATCGGACCTGGCCAGCGAACTGCTAGTTG 180

QY 2405 GGGAAATTTTCCACCATCAAAATCGCAGACTTTGGCATGAGCCGGAACCTCTATCTGGGG 2464  
DB 181 GGGAAATTTTCCACCATCAAAATCGCAGACTTTGGCATGAGCCGGAACCTCTATCTGGGG 240

QY 2465 ACTATTACCGTGTCCAGGGCGGGCGAGTGTGCTCCATCCCTGGATGCCCTGGAGTGA 2524  
DB 241 ACTATTACCGTGTCCAGGGCGGGCGAGTGTGCTCCATCCCTGGATGCCCTGGAGTGA 300



```
/map="13q12"
/clone_lib="IMAGE:1877867"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/notes="Organ: mixed (see below); Vector: pT73D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NbHM, pregnant uterus
NbHPU, and fetal heart NbH19W) were mixed, and ss circles
were made in vitro. Following HAP purification, thus DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."
BASE COUNT      122 a 112 c 108 g 104 t
ORIGIN
Query Match      12.2%; Score 444.4; DB 44; Length 446;
Best Local Similarity 99.8%; Pred. No. 1e-78; Indels 0; Gaps 0;
Matches 445; Conservative 0; Mismatches 1;
Qy 3150 CACACACTGGACCCACTGGCTGAGATCTGGGGGTGAGGAGACAAAGAGAGGAA 3209
Db 446 CACACACTGGACCCACTGGCTGAGATCTGGGGGTGAGGAGACAAAGAGAGGAA 387
Qy 3210 ATGTTTCTTGCTGCTGCTGTACTGTCTTCTTCTTCTTCTTCTTCTTCTTCT 3269
Db 386 ATGTTTCTTGCTGCTGCTGTACTGTCTTCTTCTTCTTCTTCTTCTTCTTCT 327
Qy 3270 CCTGAACACTGGACCTGGGGGTGAGCCGCCGCCAGCCCTCAGTCACCCCACTTCCCAC 3329
Db 326 CCTGAACACTGGACCTGGGGGTGAGCCGCCGCCAGCCCTCAGTCACCCCACTTCCCAC 267
Qy 3330 CTGCAGCTTCTAGCTAGAACTTCTTAAGCCTATACGTTTCTGTGGAGTAATATGGG 3389
Db 266 TTGCAGCTTCTAGCTAGAACTTCTTAAGCCTATACGTTTCTGTGGAGTAATATGGG 207
Qy 3390 ATTGGGGGGAAGAGGAGGACGACGCGCCATAGCCTTGGGGTGGACATCTCTAGTGTAGC 3449
Db 206 ATTGGGGGGAAGAGGAGGACGACGCGCCATAGCCTTGGGGTGGACATCTCTAGTGTAGC 147
Qy 3450 TGCCACATGATTTTCTATAATCACTTGGGGTGTGTACATTTTGGGGGAGAGACACA 3509
Db 146 TGCCACATGATTTTCTATAATCACTTGGGGTGTGTACATTTTGGGGGAGAGACACA 87
Qy 3510 GATTTTACACTAATATATGACCTAGCTTGGAGCAATTTTAATCCCTGCCTAGGCAG 3569
Db 86 GATTTTACACTAATATATGACCTAGCTTGGAGCAATTTTAATCCCTGCCTAGGCAG 27
Qy 3570 GTAATAAAGGTTGAGTTTCCAC 3595
Db 26 GTAATAAAGGTTGAGTTTCCAC 1
RESULT 11
LOCUS      AA831559/c
DEFINITION oc84d05.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1956393 3',
mRNA sequence.
ACCESSION  AA831559
NID         g2904658
VERSION     AA831559.1 GI:2904658
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
            Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 453)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
```

```
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL    Tumor Gene Index
COMMENT     Unpublished (1997)
            On Jan 19, 1998 this sequence version replaced gi:2286731.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 824 Std Error: 0.00
Seg primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 439.
FEATURES
    source
    1..453
    Location/Qualifiers
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:1956393"
        /clone_lib="NCI_CGAP_GCB1"
        /tissue_type="germinal center B cell"
        /lab_host="DH10B"
    /note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CBER). cDNA synthesis was
primed with a Not I-oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGGCGCCCTCATTTTTTTTTTTTTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT      122 a 111 c 108 g 112 t
ORIGIN
Query Match      12.1%; Score 440.4; DB 39; Length 453;
Best Local Similarity 99.6%; Pred. No. 6.5e-78; Indels 1; Gaps 1;
Matches 452; Conservative 0; Mismatches 1;
Qy 3149 TCACACACTGGACCCCACTGGCTGAGAACTCTGGGGTGGAGGACAAAGAGAGGAA 3208
Db 453 TCACACACTGGACCCCACTGGCTGAGAACTCTGGGGTGGAGGACAAAGAGAGGAA 394
Qy 3209 AATGTTTCTTGCTGCTGCTGTCTGTACTGTCTCAGCTTGGGCTTCTTCTCTCCATC 3268
Db 393 AATGTTTCTTGCTGCTGCTGTCTGTACTGTCTCAGCTTGGGCTTCTTCTCTCCATC 334
Qy 3269 ACCTGAACACTGGACCTGGGGGTAGCCCGCCAGCCCTCAGTCACCCCACTTCCCA 3328
Db 333 ACCTGAACACTGGACCTGGGGGTAGCCCGCCAGCCCTCAGTCACCCCACTTCCCA 275
Qy 3329 CCTGCAGCTTCTGTAGCTAGAACTTCTTAAGCCTATACGTTTCTGTGGAGTAATATGG 3388
Db 274 CCTGCAGCTTCTGTAGCTAGAACTTCTTAAGCCTATACGTTTCTGTGGAGTAATATGG 215
Qy 3389 GATTTGGGGGGAAGAGGAGGACGACGCGCCATAGCCTTGGGGTGGACATCTCTAGTGTAG 3448
Db 214 GATTTGGGGGGAAGAGGAGGACGACGCGCCATAGCCTTGGGGTGGACATCTCTAGTGTAG 155
Qy 3449 CTGCACACTGATTTTCTATAATCACTTGGGGTGTGTACATTTTGGGGGAGAGACAC 3508
Db 154 CTGCACACTGATTTTCTATAATCACTTGGGGTGTGTACATTTTGGGGGAGAGACAC 95
```



QY 3509 AGATTTTACACAAATATATATGACCTAGCTGAGGCAATTTTAAATCCCTGCACTAGGCA 3568  
 |||||  
 Db 94 AGATTTTACACAAATATATATGACCTAGCTGAGGCAATTTTAAATCCCTGCACTAGGCA 35  
 |||||  
 QY 3569 GGTAATAATAAGGTTGAGTTTCCACAAAAA 3602  
 |||||  
 Db 34 GGTAATAATAAGGTTGAGTTTCCAAAAA 1  
 |||||

RESULT 12  
 AA944928/c 542 bp mRNA EST 16-JUN-1998  
 LOCUS  
 DEFINITION EST000427 Normalized rat embryo, Bento Soares Rattus sp. cDNA clone  
 REMA085 3' end, mRNA sequence.  
 AA944928  
 ACCESSION  
 NID 93104844  
 VERSION AA944928.1 GI:3104844  
 KEYWORDS EST.  
 SOURCE Rattus sp.  
 ORGANISM Rattus sp.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 REFERENCE 1 (bases 1 to 542)  
 AUTHORS Lee.N.H., Glodek.A., Chandra.I., Mason.T.M., Quackenbush,J.,  
 Kerlavage,A.R. and Adams,M.D.  
 TITLE Rat Genome Project: Generation of a Rat EST (REST) Catalog & Rat  
 Gene Index  
 JOURNAL Unpublished (1998)  
 COMMENT On Jan 17, 1998 this sequence version replaced gi:1900074.

Contact: Lee, NH  
 ATCC  
 The Institute for Genomic Research  
 9712, Medical Center Drive, Rockville, MD 20850, USA  
 Tel: (301)-838-3529  
 Fax: (301)-838-0208  
 Email: nhlee@igir.org  
 Seq primer: M13-21.  
 Location/Qualifiers  
 1..542  
 /organism="Rattus sp."  
 /db\_xref="ATCC (inhost):2014392"  
 /db\_xref="taxon:10118"  
 /map="x"  
 /clone="REMA085"  
 /dev\_stage="embryo 8, 12, 18 dpc"  
 /note="Vector: p7T3Pac; Site\_1: EcoRI; Site\_2: NotI"  
 BASE COUNT 110 a 169 c 151 g 111 t 1 others  
 ORIGIN

Query Match 11.9%; Score 432.2; DB 40; Length 542;  
 Best Local Similarity 87.3%; Pred. No. 2.9e-76;  
 Matches 473; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 2221 AGTGCCACAGCTGAGGACAGGACGCGAGGGGGCCCTGGGGAGCGGCGAGCTGCG 2280  
 |||||  
 Db 542 AGTGCCACAGCTGAGGACAGGACGCTACTCAGGGCTCTCTGGNGACAGAGCTGAC 483  
 |||||  
 QY 2281 CAGGGGGCCACCATGACCTACCATGCTGCTGATGTGGCAGCCAGATCGCTCGCG 2340  
 |||||  
 Db 482 CAGGGGGCCACCATGACCTACCATGCTGCTGATGTGGCAGCCAGATCGCTCGCG 423  
 |||||  
 QY 2341 ATGCGCTATCTGGCCACACTCACTTTGTACATCGGACCTGGCCACCGGACCTGCTA 2400  
 |||||  
 Db 422 ATGCGATATCTGGCCACCTTAACTTTGGCATCGGGACCTGGCCACCGAACTGCTG 363  
 |||||  
 QY 2401 GTTGGGAAAATTTACCATCAAAATCGCAGACTTTGGCATGAGCGGAACTCTATGCT 2460  
 |||||  
 Db 362 GTTGGGAAAATTTACCATCAAAATCGCAGACTTTGGCATGAGCGGAACTCTACGCT 303  
 |||||

QY 2461 GGGGACTATTACCGTGTGCGAGGGCGGCGAGTCTGCCCATCGCTGATGGCTGGGAG 2520  
 |||||  
 Db 302 GGGGATTATTACCGTGTACAGGGCGGCGAGTCTGCCCATCGCTGATGGCTGGGAG 243  
 |||||  
 QY 2521 TGCATCTCATGGGGAAGTTTACGACTGCGAGTGACGTGGGGCTTTGGTGTGACCTG 2580  
 |||||  
 Db 242 TGTATCTCATGGGGAAGTTTCAACAGCCAGTGTGTGGGCTTCGGAGTGACCTG 183  
 |||||  
 QY 2581 TGGGAGGTGTGATGCTGTAGGGCCCGAGCCCTTTGGGAGCTCACCGAGCAGGCTC 2640  
 |||||  
 Db 182 TGGGAGGTGTGATGCTGTAGGGCCCGAGCCCTTTGGGAGCTCACAGAGCAGGCTC 123  
 |||||  
 QY 2641 ATCAGAACCGGGGGAGTTCTTCCGGGACAGCGCGGAGGTGTACCTGTCCCGGCGG 2700  
 |||||  
 Db 122 ATCAGAACCGGGGAGTTCTTCCAGGAGCAGCGCGGAGGTATATCTATCCAGGCCA 63  
 |||||  
 QY 2701 CCTGCTTCCCGGAGGCGCTATATGAGCTGATGCTTGGTGTGGAGCGGGAGTGTGAG 2760  
 |||||  
 Db 62 CCCGCTTCCCGGAGGCGCTTGTATGAGCTGATGCTCGGTGTGGAGCGGGAGCGGAG 3  
 |||||

QY 2761 CA 2762  
 ||  
 Db 2 CA 1  
 ||

RESULT 13  
 AA430320 429 bp mRNA EST 16-OCT-1997  
 LOCUS  
 DEFINITION zW07fill.r1 Soares\_NhMHPu\_S1 Homo sapiens cDNA clone IMAGE:768621  
 5', mRNA sequence.  
 AA430320  
 ACCESSION  
 NID 92110911  
 VERSION AA430320.1 GI:2110911  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 429)  
 AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,  
 Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,  
 Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,  
 White,Y., Wylie,T., Waterston,R. and Wilson,R.  
 WashU-Merck EST Project 1997  
 TITLE Unpublished (1997)  
 JOURNAL  
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1392778.

Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available royalty-free through LLNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Seq primer: -28ml3 rev2 ET from Amersham  
 High quality sequence stop: 411.  
 Location/Qualifiers

FEATURES  
 source

1..429  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:768621"  
 /clone\_lib="Soares\_NhMHPu\_S1"  
 /tissue\_type="Pooled human melanocyte, fetal heart, and  
 pregnant uterus"  
 /lab\_host="DH10B"  
 /note="Organ: mixed (see below); Vector: pT73D-Pac  
 (Pharmacia) with a modified polylinker; Site 1: Not I;  
 Site 2: Eco RI; Equal amounts of plasmid DNA from three  
 normalized libraries (melanocyte 2NBHM, pregnant uterus  
 NDHPU, and fetal heart NDH19W) were mixed, and ss circles  
 were made in vitro. Following HAP purification, this DNA  
 was used as tracer in a subtractive hybridization



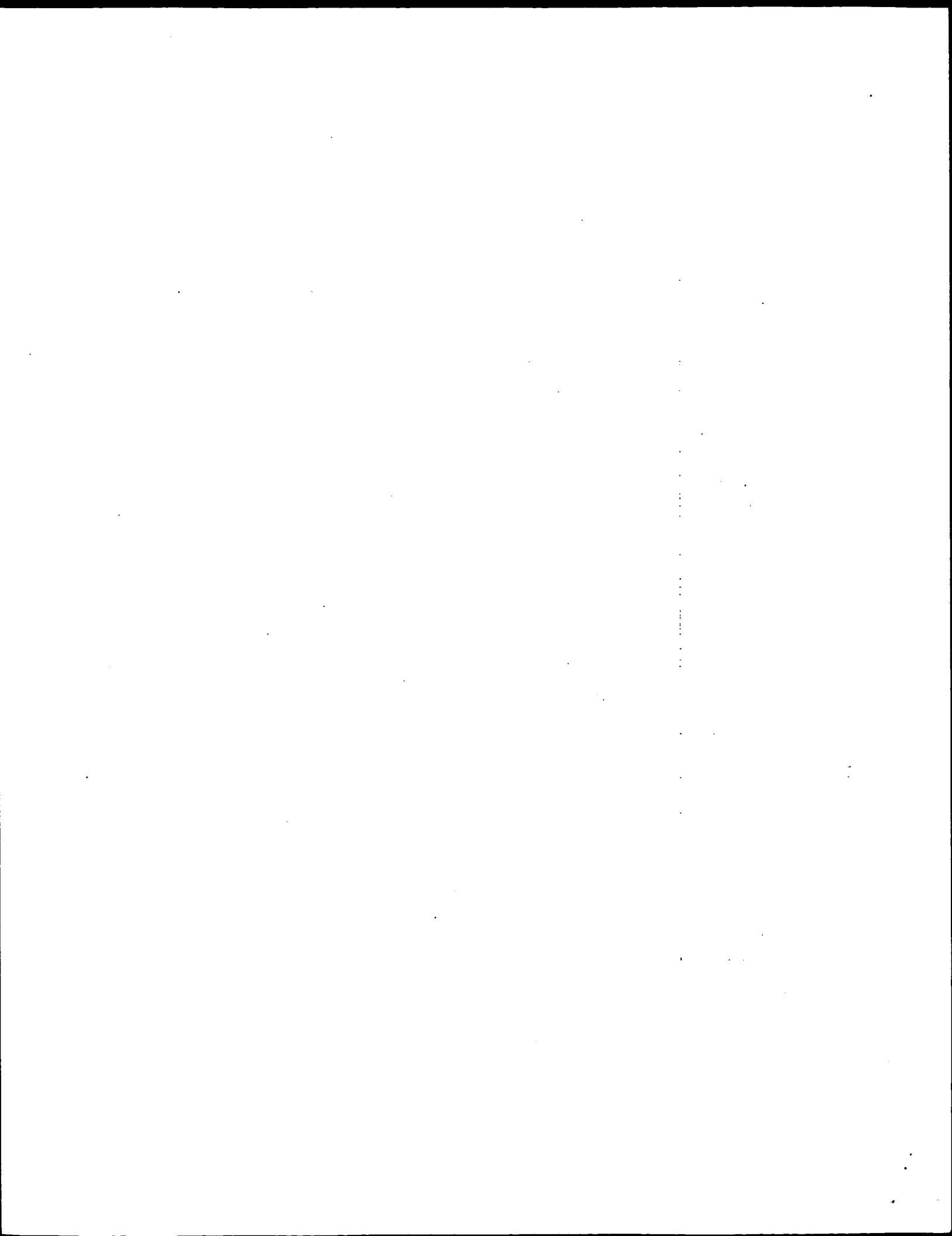
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 465)  
AUTHORS Tanaka,T., Ogiwara,A., Uchiyama,I., Takagi,T., Yazaki,Y. and Nakamura,Y.  
TITLE Construction of a Normalized Directionally Cloned cDNA Library from Adult Heart and Analysis of 3040 Clones by Partial Sequencing  
JOURNAL Genomics 35, 231-235 (1996)  
COMMENT On Nov 22, 1995 this sequence version replaced gi:1071057.

Contact: Yusuke Nakamura  
Institute of Medical Science  
University of Tokyo  
4-6-1, Shirokanedai, Minato-ku, Tokyo 108, Japan  
Tel: 81-3-5449-5372  
Fax: 81-3-5449-5433  
Email: yusuke@ims.u-tokyo.ac.jp.

FEATURES  
source  
1..465  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="3NHC2547"  
/clone\_lib="Human heart cDNA (Y Nakamura)"  
/dev\_stage="adult"  
/note="Organ: heart; normalized directionally cloned cDNA from adult heart"  
BASE COUNT 88 a 146 c 111 g 114 t 6 others  
ORIGIN

Query Match 11.4%; Score 416; DB 27; Length 465;  
Best Local Similarity 98.4%; Pred. No. 4.5e-73;  
Matches 427; Conservative 0; Mismatches 6; Indels 1; Gaps 1;  
QY 2959 AGAAGCCCTGTGCGCCACCCAGCTGGTCTGTGGATGGGATCCTCTCCACCCCTCCTCTA 3018  
Db 32 AGAAGCCCTGTGCGCCACCCAGCTGGTCTGTGGATGGGATCCTCTCCACCCCTCCTCTA 91  
QY 3019 GCCATCCCTTGGGAGGTTGGGAGAAATATAGGATAGACACTGGACATGCCCATTTGG 3078  
Db 92 GCCATCCCTTGGGAGGTTGGGAGAAATATAGGATAGACACTGGACATGCCCATTTGG 151  
QY 3079 AGCACCCTGGGCCCTGACACACTGGACAACTGATTCCTGGAGAGTGGCTGCGCCCTCAGCTTC 3138  
Db 152 AGCACCCTGGGCCCTGACACACTGGACAACTGATTCCTGGAGAGTGGCTGCGCCCTCAGCTTC 211  
QY 3139 TCTCTCCCTGTACACACTGGACCCACCTGGCTGAGATCTGGGGGTGAGGAGACAAGA 3198  
Db 212 TCTCTCCCTGTACACACTGGACCCACCTGGCTGAGATCTGGGGGTGAGGAGACAAGA 271  
QY 3199 AGGAGAGGAAATGTTCTTGTGCTGCTCTGTACTTGTCTCAGCTTGGGCTTCTTC 3258  
Db 272 AGGAGAGGAAATGTTCTTGTGCTGCTCTGTACTTGTCTCAGCTTGGGCTTCTTC 330  
QY 3259 CTCTCCATCACCCTGAACACTGGACCTGGGGGTAGCCGCCGCCCTCAGTCACCC 3318  
Db 331 CTCTCCATCACCCTGAACACTGGACCTGGGGGTAGCCGCCGCCCTCAGTCACCC 390  
QY 3319 CCACCTCCACCTGCAGTCTTGTAGTAGAATCTCTAAGCCTATACGTTTCTGTGGAG 3378  
Db 391 CNACTTCCACCTGCAGTCTTGTAGTAGAATCTCTAAGCCTATACGTTTCTGTGGAG 450  
QY 3379 TAAATATTGGGATT 3392  
Db 451 TAAATATTGGGATT 464

Search completed: November 4, 1999, 07:05:18  
Job time: 4934 sec



GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 4, 1999, 05:44:12 ; Search time 60.16 Seconds  
(without alignments)  
6289.990 Million cell updates/sec

Title: US-08-170-558-3  
Perfect score: 3637  
Sequence: 1 GAATTCGAGTCGACGCTTG.....AAAAAAAAAAAAAAAAAAAA 3637

Scoring table: IDENTITY\_NUC

Searched: 192659 seqs, 52021692 residues

Database : Issued\_Patents\_NA.\*  
1: /cgn2.6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2.6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2.6/ptodata/1/ina/5C\_COMB.seq.\*  
4: /cgn2.6/ptodata/1/ina/5D\_COMB.seq.\*  
5: /cgn2.6/ptodata/1/ina/PCUS9\_COMB.seq.\*  
6: /cgn2.6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3637	100.0	3637	2	US-08-445-640-3	Sequence 3, Appli
2	3451	94.9	3962	2	US-08-336-343A-1	Sequence 1, Appli
3	1197	32.9	1197	2	US-08-445-640-7	Sequence 7, Appli
4	669.4	18.4	3120	3	US-08-456-647B-19	Sequence 19, Appli
5	669.4	18.4	3120	3	US-08-237-401A-19	Sequence 19, Appli
6	666.8	18.3	3157	2	US-08-336-343A-3	Sequence 3, Appli
7	666.8	18.3	3157	2	US-08-336-343A-5	Sequence 5, Appli
8	195.2	5.4	2820	3	US-08-286-305A-4	Sequence 4, Appli
9	195.2	5.4	2820	4	US-08-441-104A-4	Sequence 4, Appli
10	195.2	5.4	2820	4	US-08-440-816A-4	Sequence 4, Appli
11	193.6	5.3	2301	2	US-08-306-691B-23	Sequence 23, Appli
12	193.6	5.3	2301	5	PCT-US93-06251-78	Sequence 78, Appli
13	190	5.2	3707	1	US-08-271-454-1	Sequence 1, Appli
14	190	5.2	3060	3	US-08-286-305A-6	Sequence 6, Appli
15	190	5.2	3194	3	US-08-359-705B-1	Sequence 1, Appli
16	190	5.2	3194	4	US-08-286-846A-1	Sequence 1, Appli
17	190	5.2	3060	4	US-08-441-104A-6	Sequence 6, Appli
18	190	5.2	3194	4	US-08-457-880A-1	Sequence 1, Appli
19	190	5.2	3060	4	US-08-440-816A-6	Sequence 6, Appli
20	190	5.2	3707	5	PCT-US95-08180-1	Sequence 1, Appli
21	185.2	5.1	2940	3	US-08-286-305A-8	Sequence 8, Appli
22	185.2	5.1	2940	4	US-08-441-104A-8	Sequence 8, Appli
23	185.2	5.1	2940	4	US-08-440-816A-8	Sequence 8, Appli
24	182	5.0	2526	1	US-07-912-952-1	Sequence 1, Appli
25	172.4	4.7	2463	1	US-08-339-578-1	Sequence 1, Appli
26	172.4	4.7	2463	1	US-08-339-578-1	Sequence 1, Appli
27	156.4	4.3	3398	5	PCT-US95-08493-12	Sequence 12, Appli
28	144.4	4.0	2208	5	US-08-469-537A-106	Sequence 106, App
29	144.4	4.0	2580	5	PCT-US95-08493-1	Sequence 1, Appli
30	144.4	4.0	2580	5	PCT-US95-08493-18	Sequence 18, Appli
31	144.4	4.0	2604	5	PCT-US95-08493-20	Sequence 20, Appli
32	140.6	3.9	2376	1	US-07-912-952-3	Sequence 3, Appli
33	140.6	3.9	2869	3	US-08-374-834-2	Sequence 2, Appli
34	135.2	3.7	4149	4	US-08-644-271-2	Sequence 2, Appli
35	133.2	3.7	2715	3	US-08-737-715-1	Sequence 1, Appli
36	133.2	3.7	2715	4	US-08-359-705B-5	Sequence 5, Appli
37	133.2	3.7	2715	4	US-08-286-846A-5	Sequence 5, Appli

38 131.6 3.6 2595 3 US-08-469-537A-77 Sequence 77, Appli  
39 131.6 3.6 3358 3 US-08-469-537A-104 Sequence 104, App  
40 129.6 3.6 2610 1 US-08-374-834-17 Sequence 17, Appl  
41 129.6 3.6 2610 3 US-08-644-271-28 Sequence 28, Appl  
42 121 3.3 755 3 US-08-469-537A-79 Sequence 79, Appl  
43 112.8 3.1 4989 4 US-08-666-392A-3 Sequence 3, Appli  
44 112.8 3.1 4975 4 US-08-249-687C-1 Sequence 1, Appli  
45 112.8 3.1 4989 4 US-08-625-819-1 Sequence 1, Appli

## ALIGNMENTS

RESULT 1  
US-08-445-640-3  
; Sequence 3, Application US/08445640  
; Patent No. 5709858  
; GENERAL INFORMATION:  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Mark, Melanie R.  
; APPLICANT: Scadden, David T.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Baron, Will F.  
; TITLE OF INVENTION: Protein Tyrosine Kinases  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/445,640  
; FILING DATE: 22-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/170558  
; FILING DATE: 20-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/157563  
; FILING DATE: 23-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hasak, Janet E.  
; REGISTRATION NUMBER: 28,616  
; REFERENCE/DOCKET NUMBER: 854C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-1896  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3637 bases  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-445-640-3

Query Match 100.0%; Score 3637; DB 2; Length 3637;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GAATTCGAGTCGACCTTGAGGAAATGCAAGAGATGCTGCCCCACCCCTT 60  
Db 1 GAATTCGAGTCGACCTTGAGGAAATGCAAGAGATGCTGCCCCACCCCTT 60  
QY 61 AGGCCGAGGATCAGGAGCTATGGGACCAGAGCCCTGTCTTACTGCTGCTC 120

Db 61 AGGCCCCAGGATCAGAGCTATGGACAGAGGCCCTGTCATCTTACTGTGCTGCTC 120  
Qy 121 TTGTTGCAAGTGGAGATGCTGACATGAAGGACATTTTGTATCCTGCCAAGTGGCCGCTAT 180  
Db 121 TTGTTGCAAGTGGAGATGCTGACATGAAGGACATTTTGTATCCTGCCAAGTGGCCGCTAT 180  
Qy 181 GCGCTGGGATGACAGGACCGGACCATCCACAGAGTGCATCTCTCTCTCCAGCTCCTGG 240  
Db 181 GCGCTGGGATGACAGGACCGGACCATCCACAGAGTGCATCTCTCTCTCCAGCTCCTGG 240  
Qy 241 TCAGATTCCACTGCGGCCCCACAGCAGGTGGAGAGCAGTACAGGGATGGGGCTGG 300  
Db 241 TCAGATTCCACTGCGGCCCCACAGCAGGTGGAGAGCAGTACAGGGATGGGGCTGG 300  
Qy 301 TGCCCCAGGCTGGGTGTTTCCAGAGGAGGAGTACTTTGAGGTGATCTACAAACA 360  
Db 301 TGCCCCAGGCTGGGTGTTTCCAGAGGAGGAGTACTTTGAGGTGATCTACAAACA 360  
Qy 361 CTGACCTGGTGGCTCTGGTGGGACCCAGGAGCGCATGCCGGGGCTGGGCAAGAG 420  
Db 361 CTGACCTGGTGGCTCTGGTGGGACCCAGGAGCGCATGCCGGGGCTGGGCAAGAG 420  
Qy 421 TTCTCCGGAGCTACCGGCTGGTTACTCCCGGATGGTCCCGCTGGATGGGCTGGAAG 480  
Db 421 TTCTCCGGAGCTACCGGCTGGTTACTCCCGGATGGTCCCGCTGGATGGGCTGGAAG 480  
Qy 481 GACCGCTGGGTGAGGAGTGATCTCAGGCAATGAGACCTCAGGAGTGGTGTGAAG 540  
Db 481 GACCGCTGGGTGAGGAGTGATCTCAGGCAATGAGACCTCAGGAGTGGTGTGAAG 540  
Qy 541 GACCTTGGGCCCCCATGGTTGGCCGACTGTTTCGCTTACCCCCGGGCTGACCGGGTC 600  
Db 541 GACCTTGGGCCCCCATGGTTGGCCGACTGTTTCGCTTACCCCCGGGCTGACCGGGTC 600  
Qy 601 ATGAGCTGCTGTGCGGGTAGAGCTCTATGGCTGGCTCTGGAGGATGGACTTCTGTCT 560  
Db 601 ATGAGCTGCTGTGCGGGTAGAGCTCTATGGCTGGCTCTGGAGGATGGACTTCTGTCT 560  
Qy 661 TACACCGCCCTGTGGGGCAGACAATGATTTATCTGAGGCCGTGACCTCAACGACTCC 720  
Db 661 TACACCGCCCTGTGGGGCAGACAATGATTTATCTGAGGCCGTGACCTCAACGACTCC 720  
Qy 721 ACCTATGACGACATACCGTGGGCGGACTCAGTATGGGGTCTGGGCCAGCTGGCAGAT 780  
Db 721 ACCTATGACGACATACCGTGGGCGGACTCAGTATGGGGTCTGGGCCAGCTGGCAGAT 780  
Qy 781 GGTGTGGGGCTGGATGACTTTAGGAAGTCAAGAGCTGCGGGTCTGGGCCAGGCTAT 840  
Db 781 GGTGTGGGGCTGGATGACTTTAGGAAGTCAAGAGCTGCGGGTCTGGGCCAGGCTAT 840  
Qy 841 GACTATGTGGATGGAGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGGATTTGAG 900  
Db 841 GACTATGTGGATGGAGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGGATTTGAG 900  
Qy 901 TTTGACCGGCTGAGGGCTTCCAGGCTATCAGTCCACTGTAAACAATGCACACGCTG 960  
Db 901 TTTGACCGGCTGAGGGCTTCCAGGCTATCAGTCCACTGTAAACAATGCACACGCTG 960  
Qy 961 GGACCCGCTGCTGCGGGGTGGAAATGCTGCTTCCGGGCTGGCCCTGGCATGGGCTGG 1020  
Db 961 GGACCCGCTGCTGCGGGGTGGAAATGCTGCTTCCGGGCTGGCCCTGGCATGGGCTGG 1020  
Qy 1021 GAGGGGAGCCCATGCGCCACAACCTAGGGGCAACCTGGGGACCCACAGAGCCGGGCT 1080  
Db 1021 GAGGGGAGCCCATGCGCCACAACCTAGGGGCAACCTGGGGACCCACAGAGCCGGGCT 1080  
Qy 1081 GTCTAGTGGCCCTTGGGCGGCTGGCTCTTCTGAGTGGCCCTTCTCTTTTGG 1140  
Db 1081 GTCTAGTGGCCCTTGGGCGGCTGGCTCTTCTGAGTGGCCCTTCTCTTTTGG 1140  
Qy 1141 GGGCCCTGGTCTTTCAGGGAATCTCTTCTATCTCTGATGTGGTGAACAATTTCTCT 1200  
Db 1141 GGGCCCTGGTCTTTCAGGGAATCTCTTCTATCTCTGATGTGGTGAACAATTTCTCT 1200

Db 1141 GGGCCCTGGTCTTTCAGGGAATCTCTTCTATCTCTGATGTGGTGAACAATTTCTCT 1200  
Qy 1201 CCGGCACTGGAGGACACTTCCCGCCAGCCCCCTGGTGGCGGCTGCCACCTCCCAAC 1260  
Db 1201 CCGGCACTGGAGGACACTTCCCGCCAGCCCCCTGGTGGCGGCTGCCACCTCCCAAC 1260  
Qy 1261 AACTTCAGCAGCTTGGAGCTGGAGCCAGAGGCCAGAGCCCGTGGCCAAAGCCGAGGG 1320  
Db 1261 AACTTCAGCAGCTTGGAGCTGGAGCCAGAGGCCAGAGCCCGTGGCCAAAGCCGAGGG 1320  
Qy 1321 AGCCGACCGCATCTCATCGGCTGCTGTGGCCATCATCTCTCTCTCTCTCTCTCT 1380  
Db 1321 AGCCGACCGCATCTCATCGGCTGCTGTGGCCATCATCTCTCTCTCTCTCTCTCT 1380  
Qy 1381 ATTGCGCTCATCTCTGCGGCTGCACTGCGCAGGCTCTCAGCAAGGCTGAACGAGG 1440  
Db 1381 ATTGCGCTCATCTCTGCGGCTGCACTGCGCAGGCTCTCAGCAAGGCTGAACGAGG 1440  
Qy 1441 GTGTTGGAAGAGGAGCTGACGGTTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1500  
Db 1441 GTGTTGGAAGAGGAGCTGACGGTTTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1500  
Qy 1501 AACCGCCAGCTCTAGAGAGCCACCCCTACAGAGCCCGGCTCTGGGAATCCG 1560  
Db 1501 AACCGCCAGCTCTAGAGAGCCACCCCTACAGAGCCCGGCTCTGGGAATCCG 1560  
Qy 1561 CCCCACTCCGCTCTCTGTCGCCCAATGGCTCTGCTCTCTCTCTCTCTCTCTCTCTCT 1620  
Db 1561 CCCCACTCCGCTCTCTGTCGCCCAATGGCTCTGCTCTCTCTCTCTCTCTCTCTCTCT 1620  
Qy 1621 CGCTCTCTCTGCGGCACTTACGCGCGTCCCTCTGAGGCGCGGCGGCGGCGGCGGCG 1680  
Db 1621 CGCTCTCTCTGCGGCACTTACGCGCGTCCCTCTGAGGCGCGGCGGCGGCGGCGGCG 1680  
Qy 1681 TGGGCAAAACCCACACAGCCAGGCTACAGTGGGCACTATATGGAGCTGAGAAGCCA 1740  
Db 1681 TGGGCAAAACCCACACAGCCAGGCTACAGTGGGCACTATATGGAGCTGAGAAGCCA 1740  
Qy 1741 GCGCGCCGCTCTTCTGCGGCGGCACTTACGCGCGTCCCTCTGAGGCGCGGCGGCGGCG 1800  
Db 1741 GCGCGCCGCTCTTCTGCGGCGGCACTTACGCGCGTCCCTCTGAGGCGCGGCGGCGGCG 1800  
Qy 1801 ATTGTTACCTTGCAGGCGCTCACCGGGGCAACACCTATGCTGTGCTGCACTGCCGCG 1860  
Db 1801 ATTGTTACCTTGCAGGCGCTCACCGGGGCAACACCTATGCTGTGCTGCACTGCCGCG 1860  
Qy 1861 GGGCAGCTCGGGATGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1920  
Db 1861 GGGCAGCTCGGGATGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1920  
Qy 1921 GAGAGCTTGGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1980  
Db 1921 GAGAGCTTGGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1980  
Qy 1981 GATCTGTGCTGATTTTCCGCTTAAATGTGGTAAAGGACACCTTTTGTGGTGGT 2040  
Db 1981 GATCTGTGCTGATTTTCCGCTTAAATGTGGTAAAGGACACCTTTTGTGGTGGT 2040  
Qy 2041 GTCAGATCTTACGGCAGATGCCACCAAGATGCCAGGAATGATTTCTTGAAGAGG 2100  
Db 2041 GTCAGATCTTACGGCAGATGCCACCAAGATGCCAGGAATGATTTCTTGAAGAGG 2100  
Qy 2101 AAGATCATGTGAGGCTCAAGGACCCCAACATCATTTGGGCTGCTGGGCGTGTGTGCG 2160  
Db 2101 AAGATCATGTGAGGCTCAAGGACCCCAACATCATTTGGGCTGCTGGGCGTGTGTGCG 2160  
Qy 2161 GACGACCCCTTCTGATGATTTACTGACTACATGAGAACGCGGACCTCAACCACTTCT 2220  
Db 2161 GACGACCCCTTCTGATGATTTACTGACTACATGAGAACGCGGACCTCAACCACTTCT 2220  
Qy 2221 AGTGGCCACAGCTGGAGGCAAGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2280  
Db 2221 AGTGGCCACAGCTGGAGGCAAGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2280

QY 2281 CAGGGGCCACCATCAGCTACCAATGCTCTGCTGATGTGGCAGCCAGATCGCCTCCGGC 2340  
DB 2281 CAGGGGCCACCATCAGCTACCAATGCTCTGCTGATGTGGCAGCCAGATCGCCTCCGGC 2340  
QY 2341 ATGGCTATCTGGCCACACCTCACTTTGTATCATCGGGACCTGGCCACCGGGAACCTGCCTA 2400  
DB 2341 ATGGCTATCTGGCCACACCTCACTTTGTATCATCGGGACCTGGCCACCGGGAACCTGCCTA 2400  
QY 2401 GTTGGGAAATTTACCATCAAAATCGCAGACTTTGGCATGAGCCGGAACCTCTATGCT 2460  
DB 2401 GTTGGGAAATTTACCATCAAAATCGCAGACTTTGGCATGAGCCGGAACCTCTATGCT 2460  
QY 2461 GGGGACTATTACCGTGTGAGGGCGGGGAGTCTGCTGCCATCCGCTGGATGGGAG 2520  
DB 2461 GGGGACTATTACCGTGTGAGGGCGGGGAGTCTGCTGCCATCCGCTGGATGGGAG 2520  
QY 2521 TGCATCTCATGGGGAAGTTACGACTGGGAGTGAGTGTGGCCCTTTGGTGTGACCCCTG 2580  
DB 2521 TGCATCTCATGGGGAAGTTACGACTGGGAGTGAGTGTGGCCCTTTGGTGTGACCCCTG 2580  
QY 2581 TGGGAGTGTGATGCTCTGTAGGGCCCGAGCCCTTTGGGAGCTCACCGACGAGGTC 2640  
DB 2581 TGGGAGTGTGATGCTCTGTAGGGCCCGAGCCCTTTGGGAGCTCACCGACGAGGTC 2640  
QY 2641 ATCGAGAACCGGGGAGTCTTCCGGGACCGAGGCGCGGAGTGTACCTGTCCCGGCGG 2700  
DB 2641 ATCGAGAACCGGGGAGTCTTCCGGGACCGAGGCGCGGAGTGTACCTGTCCCGGCGG 2700  
QY 2701 CTTGCTGCTCCCGGAGGCTTATAGCTGATGCTTCCGCTGGGAGGCTGAGTGTGAG 2760  
DB 2701 CTTGCTGCTCCCGGAGGCTTATAGCTGATGCTTCCGCTGGGAGGCTGAGTGTGAG 2760  
QY 2761 CAGGACACCCCTTTTCCAGCTGATGCTGCTTCCGAGAGGATGCTCAACACGCTG 2820  
DB 2761 CAGGACACCCCTTTTCCAGCTGATGCTGCTTCCGAGAGGATGCTCAACACGCTG 2820  
QY 2821 TGAATCACATCCAGCTGCTGCTTCCGAGGAGTGTATCCAGGAGGATGCTCAACACGCT 2880  
DB 2821 TGAATCACATCCAGCTGCTGCTTCCGAGGAGTGTATCCAGGAGGATGCTCAACACGCT 2880  
QY 2881 AAAAAGAGGACACATGCGACCTCTGCTTCCGCTTCCGAGGAGGATGCTCAACACGCT 2940  
DB 2881 AAAAAGAGGACACATGCGACCTCTGCTTCCGCTTCCGAGGAGGATGCTCAACACGCT 2940  
QY 2941 TAGAGCAGTGAAGTGCAGAGCCCTCTGCTCCGACGAGTGTGCTGCTGATGGAT 3000  
DB 2941 TAGAGCAGTGAAGTGCAGAGCCCTCTGCTCCGACGAGTGTGCTGCTGATGGAT 3000  
QY 3001 CCTCTCCACCTCTCTAGCCATCCCTTGGGAGGAGTGGGAGGAGTATAGGATAGACA 3060  
DB 3001 CCTCTCCACCTCTCTAGCCATCCCTTGGGAGGAGTGGGAGGAGTATAGGATAGACA 3060  
QY 3061 CTGGACATGGCCCATTTGGAGCAGCTGGGCGCCACCTGGACACACATGATTCCTGGAGAGT 3120  
DB 3061 CTGGACATGGCCCATTTGGAGCAGCTGGGCGCCACCTGGACACACATGATTCCTGGAGAGT 3120  
QY 3121 GGTGGGCCCCAGCTTCTCTCTCTGTCACACACTGGACCCCACTGGCTGAGAACTCTG 3180  
DB 3121 GGTGGGCCCCAGCTTCTCTCTCTGTCACACACTGGACCCCACTGGCTGAGAACTCTG 3180  
QY 3181 GGGGTGAGGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3240  
DB 3181 GGGGTGAGGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3240  
QY 3241 CTCAGCTTGGGCTTCT 3300  
DB 3241 CTCAGCTTGGGCTTCT 3300  
QY 3301 CCCAGCCCTCAGTCAACCCCACTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3360  
DB 3301 CCCAGCCCTCAGTCAACCCCACTTCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3360

QY 3361 CTATACGTTTCTGTGGAGTAATATTGGGATTGGGGAGAGAGAGGAGGAGGAGGAGGAGGAGG 3420  
DB 3361 CTATACGTTTCTGTGGAGTAATATTGGGATTGGGGAGAGAGAGGAGGAGGAGGAGGAGGAGG 3420  
QY 3421 GCCTTGGGTTGGACATCTCTAGTGTAGTCCACATTTGATTTTCTATATATCACTTGGG 3480  
DB 3421 GCCTTGGGTTGGACATCTCTAGTGTAGTCCACATTTGATTTTCTATATCACTTGGG 3480  
QY 3481 GTTGTACATTTTGGGGGAGAGACACAGATTTTACACTAATATATGGACCTAGCTTGG 3540  
DB 3481 GTTGTACATTTTGGGGGAGAGACACAGATTTTACACTAATATATGGACCTAGCTTGG 3540  
QY 3541 AGGCAATTTTAAATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3600  
DB 3541 AGGCAATTTTAAATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3600  
QY 3601 AAAAAAATTTTAAATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3637  
DB 3601 AAAAAAATTTTAAATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3637

RESULT 2

US-08-336-343A-1  
; Sequence 1, Application US/08336343A  
; Patent No. 5677144  
; GENERAL INFORMATION:  
; APPLICANT: Ullrich, Axel  
; APPLICANT: Alves, Frauke  
; TITLE OF INVENTION: CKK-2, A No. 5677144el Receptor Tyrosine Kinase  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/336,343A  
; FILING DATE: 08-NOV-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7683-065  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-9741/8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3962 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: unknown  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 321...3077  
; US-08-336-343A-1

Query Match 94.9%; Score 3451; DB 2; Length 3962;  
Best Local Similarity 97.0%; Pred. No. 0;  
Matches 3589; Conservative 0; Mismatches 5; Indels 105; Gaps 3;

QY 17 GTTGACTTGAAGGAATGCCAAGAGATGCTGCCCCACCCTTAGCGCCGAGGATCAG 76  
DB 256 GTTGACTTGAAGGAATGCCAAGAGATGCTGCCCCACCCTTAGCGCCGAGGATCAG 315  
QY 77 GAGTATGGGACAGAGGCCCTGTCTATCTTTACTGTGCTGCTCTTGGTGGAAGTGAG 136  
DB 316 GAGTATGGGACAGAGGCCCTGTCTATCTTTACTGTGCTGCTCTTGGTGGAAGTGAG 375  
QY 137 ATGTGACATGAAGGACATTTGATCTGCTGCAAGTGCCTGATGCCCTGGGATCGAG 196  
DB 376 ATGTGACATGAAGGACATTTGATCTGCTGCAAGTGCCTGATGCCCTGGGATCGAG 435  
QY 197 ACCGACCATCCAGACAGTACATCTCTGCTCCAGTCTGCTGCTGATTCACATCCG 256  
DB 436 ACCGACCATCCAGACAGTACATCTCTGCTCCAGTCTGCTGCTGATTCACATCCG 495  
QY 257 CCCGCCACAGAGGTTGAGAGCAGTACGCGGGGATGGGCTGTGCGCCGCGAGGTCGG 316  
DB 496 CCCGCCACAGAGGTTGAGAGCAGTACGCGGGGATGGGCTGTGCGCCGCGAGGTCGG 555  
QY 317 TGTTCCTCCAGGAGGAGTACTTGCAGGTGGATCTACACGACTGCACCTGGTGCTC 376  
DB 556 TGTTCCTCCAGGAGGAGTACTTGCAGGTGGATCTACACGACTGCACCTGGTGCTC 615  
QY 377 TGGTGGGACCCAGGACGCGATGCGGGGCTGGGCAAGGATTTCCCGGAGCTACC 436  
DB 616 TGGTGGGACCCAGGACGCGATGCGGGGCTGGGCAAGGATTTCCCGGAGCTACC 675  
QY 437 GCGTCGTTACTCCGGGATGGTGCCTGATGGGCTGGAAGGACCGTGGGGTCAGG 496  
DB 676 GCGTCGTTACTCCGGGATGGTGCCTGATGGGCTGGAAGGACCGTGGGGTCAGG 735  
QY 497 AGGTGATCTCAGGCAATGAGGACCTGAGGAGTGGTCTGAAGGACCTTGGGCCCCCA 556  
DB 736 AGGTGATCTCAGGCAATGAGGACCTGAGGAGTGGTCTGAAGGACCTTGGGCCCCCA 795  
QY 557 TGGTGGCCGACTGCTTCTTACCCTGGGCTGACCGGCTCATGAGCTCTGTCTGC 616  
DB 796 TGGTGGCCGACTGCTTCTTACCCTGGGCTGACCGGCTCATGAGCTCTGTCTGC 855  
QY 617 GGGTAGAGCTCTATGGTGCCTCTGGAGGATGGTCTGCTTACACCGCCCTGTGG 676  
DB 856 GGGTAGAGCTCTATGGTGCCTCTGGAGGATGGTCTGCTTACACCGCCCTGTGG 915  
QY 677 GCGACAGATGATTTATCTGAGCGCTGTACTCAAGGACTCCACCTATGACGACATA 736  
DB 916 GCGACAGATGATTTATCTGAGCGCTGTACTCAAGGACTCCACCTATGACGACATA 975  
QY 737 CCGTGGCGGACTGCAAGTATGGGCTGTGGCCAGCTGGCAGATGGTGTGGGCTGG 796  
DB 976 CCGTGGCGGACTGCAAGTATGGGCTGTGGCCAGCTGGCAGATGGTGTGGGCTGG 1035  
QY 797 ATGACTTTAGGAAGTACAGAGCTGCGGCTCTGCCAGGCTATGACTATGTGGATGGA 856  
DB 1036 ATGACTTTAGGAAGTACAGAGCTGCGGCTCTGCCAGGCTATGACTATGTGGATGGA 1095  
QY 857 GCAACACAGCTCTCCAGTGGCTATGTGGAGATGGATTTGAGTTTACCGGCTGAGGG 916  
DB 1096 GCAACACAGCTCTCCAGTGGCTATGTGGAGATGGATTTGAGTTTACCGGCTGAGGG 1155  
QY 917 CTTTCAGGCTATGAGGTGCTACTGTAAACATGCACACGCTGGAGCCGCTGCGCTG 976  
DB 1156 CTTTCAGGCTATGAGGTGCTACTGTAAACATGCACACGCTGGAGCCGCTGCGCTG 1215  
QY 977 GCGGGTGAATGCTCTTCCGCGTGGCCCTGCCATGGCTGGGAGGGAGCCCATGC 1036  
DB 1216 GCGGGTGAATGCTCTTCCGCGTGGCCCTGCCATGGCTGGGAGGGAGCCCATGC 1275  
QY 1037 GCAACACCTAGGGGCAACCTGGGGACCCAGACCGCGGCTGTCTCAGTGCCTTGG 1096  
DB 1276 GCAACACCTAGGGGCAACCTGGGGACCCAGACCGCGGCTGTCTCAGTGCCTTGG 1335  
QY 1097 GCGCGGCTGTGCTGCTTCTGCAAGTGGGCTCTCTTTCGGGGGCTGCTTACTCT 1156

DB 1336 GCGGCGTGTGCTGCTTCTGAGTCCGCTTCTCTTTGCGGGGCTTGGTACTCT 1395  
QY 1157 TCAGGGAATCTCTCTTCTGATGTGGTGAACAATCTCTCCGCACTGGGAGCA 1216  
DB 1396 TCAGGGAATCTCTCTTCTGATGTGGTGAACAATCTCTCCGCACTGGGAGCA 1455  
QY 1217 CTTCCGCGAGCCCTGCTGGGCGCTGCGCCACCTCCCAACATTCAGCAGCTGG 1276  
DB 1456 CTTCCGCGAGCCCTGCTGGGCGCTGCGCCACCTCCCAACATTCAGCAGCTGG 1515  
QY 1277 AGCTGGAGCCAGAGCCAGAGCCGCTGGCAAGCCCGAGGGAGCCGACCCCATCC 1336  
DB 1516 AGCTGGAGCCAGAGCCAGAGCCGCTGGCAAGCCCGAGGGAGCCGACCCCATCC 1575  
QY 1337 TCATCGGCTGCTGTGGCCATCTCTGCTGCTCTCATCTTTCCTCATCTGCTCT 1396  
DB 1576 TCATCGGCTGCTGTGGCCATCTCTGCTGCTCTCATCTTTCCTCATCTGCTCT 1635  
QY 1397 GCGGCTGCACTGGGCGAGGCTCTCAGCAAGGCTGAACGAGGCTGTGGAAGAGGAGC 1456  
DB 1636 GCGGCTGCACTGGGCGAGGCTCTCAGCAAGGCTGAACGAGGCTGTGGAAGAGGAGC 1695  
QY 1457 TGACGGTTCACCTCTGCTGCTGCGGACATCTCTCATCAACACCCGCGCTCCTA 1516  
DB 1696 TGACGGTTCACCTCTGCTGCTGCGGACATCTCTCATCAACACCCGCGCTCCTA 1755  
QY 1517 GAGAGCACCCCGCTACAGAGAGCCCGGCTCTGCGGAATCCGCGCCACTCCGCTCCT 1576  
DB 1756 GAGAGCACCCCGCTACAGAGAGCCCGGCTCTGCGGAATCCGCGCCACTCCGCTCCT 1815  
QY 1577 GTGTCCCAATGGCTGTGCTGCTCTCAATCCAGCCTACCGCTCTCTTGGGCA 1636  
DB 1816 GTGTCCCAATGGCTGTGCTGCTCTCAATCCAGCCTACCGCTCTCTTGGGCA 1875  
QY 1637 CTTACGCGCTCCCTCGAGCGCGGCGCCCGCCACACCGCTGGGCAACCCACA 1696  
DB 1876 CTTACGCGCTCCCTCGAGCGCGGCGCCCGCCACACCGCTGGGCAACCCACA 1935  
QY 1697 ACACCCAGGCTACAGTGGGACTATATGAGGCTGAGAGCCAGGCGCCGCTTCTGC 1756  
DB 1936 ACACCCAGGCTACAGTGGGACTATATGAGGCTGAGAGCCAGGCGCCGCTTCTGC 1995  
QY 1757 CCGCACTCCCAAGACAGCGTCCCGCTATGCGAGGCTGACATTTACCTTCAGG 1816  
DB 1996 CCGCACTCCCAAGACAGCGTCCCGCTATGCGAGGCTGACATTTACCTTCAGG 2055  
QY 1817 GCGTACCGGGGCAACACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1876  
DB 2056 GCGTACCGGGGCAACACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2115  
QY 1877 GCGCGCCAGAGTGGATTTCCCTCGATCTCGACTCCGCTTCAAGGAGAGCTTGGCAGG 1936  
DB 2116 GCGCGCCAGAGTGGATTTCCCTCGATCTCGACTCCGCTTCAAGGAGAGCTTGGCAGG 2175  
QY 1937 GCGAGTTGGGAGTGCACCTGTGTGAGGTGACAGCCCTCAAGATCTTGGTCACTTGG 1996  
DB 2176 GCGAGTTGGGAGTGCACCTGTGTGAGGTGACAGCCCTCAAGATCTTGGTCACTTGG 2235  
QY 1997 ATTTCCCGCTTAATGTGCGTAAAGGACACCTTTGCTGCTGCTGCTGCTGCTGCTGCTG 2056  
DB 2236 ATTTCCCGCTTAATGTGCGTAAAGGACACCTTTGCTGCTGCTGCTGCTGCTGCTGCTG 2295  
QY 2057 CAGATGCCACCAAGATG-----CCAGGAATGATTTCTTGAAGAGG 2098  
DB 2296 CAGATGCCACCAAGATG-----CCAGGAATGATTTCTTGAAGAGG 2355  
QY 2099 TGAAGATCATGTGAGGCTCAAGGACCCCAACATCATCTGCGTCTGGGCTGTGTGTC 2158  
DB 2356 TGAAGATCATGTGAGGCTCAAGGACCCCAACATCATCTGCGTCTGGGCTGTGTGTC 2415  
QY 2159 AGGACGACCCCTCTGCTGATTTACTGACTTACATGGAGACGCGGACCTCAACAGTTCC 2218



Db 2416 AGGACGACCCCTCTGCTGATGATTACTGATGATGAGAACGGCGACCTCAACCACTTCC 2475  
Qy 2219 TCAGTGCCCAACAGCTGGAGACAAGCAGCCGAGGGGGCCCTGGGGAGCGGCGAGGCTG 2278  
Db 2476 TCAGTGCCCAACAGCTGGAGACAAGCAGCCGAGGGGGCCCTGGGGAGCGGCGAGGCTG 2535  
Qy 2279 CGAGGGGGCCCAACATCAGCTACCCATGCTGCTGATGATGGGAGCCGAGATCGCTCCG 2338  
Db 2536 CGAGGGGGCCCAACATCAGCTACCCATGCTGCTGATGATGGGAGCCGAGATCGCTCCG 2595  
Qy 2339 GCATGCGCTATCTGGCCACACTCAACTTTGTACATCGGACCTGGCCAGCGGAACCTGCC 2398  
Db 2596 GCATGCGCTATCTGGCCACACTCAACTTTGTACATCGGACCTGGCCAGCGGAACCTGCC 2655  
Qy 2399 TAGTTGGGAAAATTTACCATCAAAATCGACAGCTTTGGCATGAGCCGGAACCTCTATG 2458  
Db 2656 TAGTTGGGAAAATTTACCATCAAAATCGACAGCTTTGGCATGAGCCGGAACCTCTATG 2715  
Qy 2459 CTGGGGACTATTACCGTGTGACGGCCGGGAGTGTGCTGCCCATCCGCTGGATGGCTGGG 2518  
Db 2716 CTGGGGACTATTACCGTGTGACGGCCGGGAGTGTGCTGCCCATCCGCTGGATGGCTGGG 2775  
Qy 2519 AGTGCATCTCTATGGGGAAGTTTACGAGCTCGGAGTGACGFTGGGGCTTTGGTGTGACCC 2578  
Db 2776 AGTGCATCTCTATGGGGAAGTTTACGAGCTCGGAGTGACGFTGGGGCTTTGGTGTGACCC 2835  
Qy 2579 TGTGGGAGTGTGATGCTCTGTAGGCCCCAGCCCTTTGGGAGCTCACCGACGACGAG 2638  
Db 2836 TGTGGGAGTGTGATGCTCTGTAGGCCCCAGCCCTTTGGGAGCTCACCGACGACGAG 2895  
Qy 2639 TCATGAGAACGGGGGAGTGTCTCCGGGACAGGGCCGGGAGGTGATCTGTCGCCGGC 2698  
Db 2896 TCATGAGAACGGGGGAGTGTCTCCGGGACAGGGCCGGGAGGTGATCTGTCGCCGGC 2955  
Qy 2699 CGCTGCTGCTGCCGAGGGCTTATGAGCTGATGCTTGGGTCTGAGCGCGGAGTCTG 2758  
Db 2956 CGCTGCTGCTGCCGAGGGCTTATGAGCTGATGCTTGGGTCTGAGCGCGGAGTCTG 3015  
Qy 2759 AGCAGGACCAACCTTTTCCAGCTGCATCGGTTCTGCGAGAGGATGCACTCAACACGG 2818  
Db 3016 AGCAGGACCAACCTTTTCCAGCTGCATCGGTTCTGCGAGAGGATGCACTCAACACGG 3075  
Qy 2819 TGTGAATCACACATCAGCTGCCCTCCCTCAGGGAGTGATCCAGGGGAAGCCAGTGACA 2878  
Db 3076 TGTGAATCACACATCAGCTGCCCTCCCTCAGGGAGTGATCCAGGGGAAGCCAGTGACA 3135  
Qy 2879 CTAACAAGAGGACACAATGGACCTCTGCCCTTCCCTCCGAGACGCCATCACTCT 2938  
Db 3136 CTAACAAGAGGACACAATGGACCTCTGCCCTTCCCTCCGAGACGCCATCACTCT 3195  
Qy 2939 AATAGAGCAGTGAGACTGC----- 2959  
Db 3196 AATAGAGCAGTGAGACTGCAGGTGGGCTGGGCCACCCAGGAGGTGATGCCCTTCTC 3255  
Qy 2959 -----AGAGGCCCTGTGCG 2972  
Db 3256 CCCTTCTGGACACTCTCATGCTCCCTTCTCTGTTCTTCTCTCCATGAGAGCCCTGTGCG 3315  
Qy 2973 CCACCCAGCTGTCTGTGGATGGATCCTCTCCACCTCCTCTAGCCATCCCTTGGGG 3032  
Db 3316 CCACCCAGCTGTCTGTGGATGGATCCTCTCCACCTCCTCTAGCCATCCCTTGGGG 3375  
Qy 3033 AAGGGTGGGAGAAATATAGATAGACTGGACATGGCCCATTTGGAGCACCTTGGGCCCC 3092  
Db 3376 AAGGGTGGGAGAAATATAGATAGACTGGACATGGCCCATTTGGAGCACCTTGGGCCCC 3435  
Qy 3093 ACTGGACAACATGATCTCTGGAGAGGTGCTCGGCCCCAGCTTCTCTCTCCGTGTAC 3152  
Db 3436 ACTGGACAACATGATCTCTGGAGAGGTGCTCGGCCAGCTTCTCTCTCCGTGTAC 3494  
Qy 3153 ACAGTGACCCCACTGGCTGAGATCTGGGGTGGAGGAGGACAGAGGAGGAGAAATG 3212  
Db 3495 ACAGTGACCCCACTGGCTGAGATCTGGGGTGGAGGAGGACAGAGGAGGAGAAATG 3554

Qy 3213 TTTCCTTGTGCTGCTCCTGTACTTGTCTCAGCTTGGGCTTCTTCTCCTCCATCACCT 3272  
Db 3555 TTTCCTTGTGCTGCTCCTGTACTTGTCTCAGCTTGGGCTTCTTCTCCTCCATCACCT 3614  
Qy 3273 GAACACATGGACCTGGGGGTAGCCCCAGCCCCAGCTCAGTCAACCCCACTTCCCACTG 3332  
Db 3615 GAACACATGGACCTGGGGGTAGCCCCAGCCCCAGCTCAGTCAACCCCACTTCCCACTG 3674  
Qy 3333 CAGCTTCTAGCTAGAACTTCTTAAGCCTATAGCTTCTGTGGAGTAAATATTGGGATT 3392  
Db 3675 CAGCTTCTAGCTAGAACTTCTTAAGCCTATAGCTTCTGTGGAGTAAATATTGGGATT 3734  
Qy 3393 GGGGGAAGAGGAGCAACGGCCCATAGCCTTGGGGTGGACATCTCTAGTGTAGCTGC 3452  
Db 3735 GGGGGAAGAGGAGCAACGGCCCATAGCCTTGGGGTGGACATCTCTAGTGTAGCTGC 3794  
Qy 3453 CACATGATTTTCTATATACACTTGGGGTTGTACATTTTGGGGGAGAGACACAGAT 3512  
Db 3795 CACATGATTTTCTATATACACTTGGGGTTGTACATTTTGGGGGAGAGACACAGAT 3854  
Qy 3513 TTTTACACTAATATATGACCTAGCTTGGGCAATTTTAAATCCCTGCCTAGGACAGTA 3572  
Db 3855 TTTTACACTAATATATGACCTAGCTTGGGCAATTTTAAATCCCTGCCTAGGACAGTA 3914  
Qy 3573 ATAATAAGGTGTAGTTTCCACAAAAAATAAAAAA 3611  
Db 3915 ATAATAAGGTGTAGTTTCCACAAAAAATAAAAAA 3953

## RESULT 3

US-08-445-640-7  
; Sequence 7, Application US/08445640  
; Patent No. 5709858  
; GENERAL INFORMATION:  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Mark, Melanie R.  
; APPLICANT: Scadden, David T.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Baron, Will F.  
; TITLE OF INVENTION: Protein Tyrosine Kinases  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/445,640  
; FILING DATE: 22-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/170558  
; FILING DATE: 20-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/157563  
; FILING DATE: 23-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hasak, Janet E.  
; REGISTRATION NUMBER: 28,616  
; REFERENCE/DOCKET NUMBER: 854C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-1896  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 7:

## SEQUENCE CHARACTERISTICS:

LENGTH: 1197 bases  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-445-640-7

Query Match 32.9%; Score 1197; DB 2; Length 1197;

Best Local Similarity 100.0%; Pred. No. 1.3e-256;

Matches 1197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 GATGCTGACATGAGGAGACATTTTATCTCCCAAGTGGCGCTATGCGCTGGGCGATGCAG 195  
DB 1 GATGCTGACATGAGGAGACATTTTATCTCCCAAGTGGCGCTATGCGCTGGGCGATGCAG 60  
QY 196 GACCGGACATCCACAGACAGTGCATCTCTCTCCAGCTCCCTGGTGCAGATTCACATGCC 255  
DB 61 GACCGGACATCCACAGACAGTGCATCTCTCTCCAGCTCCCTGGTGCAGATTCACATGCC 120  
QY 256 GCGCCGCCACAGAGTTTGGAGAGCAGTGCACGGGATGGGCGCTGGTGGCCCGCAGGGGTGG 315  
DB 121 GCGCCGCCACAGAGTTTGGAGAGCAGTGCACGGGATGGGCGCTGGTGGCCCGCAGGGGTGG 180  
QY 316 GTGTTTCCCAAGGAGGAGTACTTTCAGAGTGGATCTACAACTGACACTGGCTGGTGGCT 375  
DB 181 GTGTTTCCCAAGGAGGAGTACTTTCAGAGTGGATCTACAACTGACACTGGCTGGTGGCT 240  
QY 376 CTGTTGGGACCCAGGAGGACGATCCCGGGGCTGGGCAAGGAGTCTCCCGGAGCTAC 435  
DB 241 CTGTTGGGACCCAGGAGGACGATCCCGGGGCTGGGCAAGGAGTCTCCCGGAGCTAC 300  
QY 436 CGGCTGCGTCTACTCCCGGATGGTGGCGCTGGATGGGCTGGAGGACCGGTGGGGTTCAG 495  
DB 301 CGGCTGCGTCTACTCCCGGATGGTGGCGCTGGATGGGCTGGAGGACCGGTGGGGTTCAG 360  
QY 496 GAGGTGATCTCAGGCAATGAGACCTGAGGAGTGGTGTGAAGACCTTTGGGCCCGCCC 555  
DB 361 GAGGTGATCTCAGGCAATGAGACCTGAGGAGTGGTGTGAAGACCTTTGGGCCCGCCC 420  
QY 556 ATGGTTGGCCGACGTTTCGCTTCTACCCCGGCTGACCGGGTCATGAGGCTGTGCTG 615  
DB 421 ATGGTTGGCCGACGTTTCGCTTCTACCCCGGCTGACCGGGTCATGAGGCTGTGCTG 480  
QY 616 CGGCTAGAGCTCTATGGCTGCTCTGGAGGATGACTCTCTTACACCGCCCTGTG 675  
DB 481 CGGCTAGAGCTCTATGGCTGCTCTGGAGGATGACTCTCTTACACCGCCCTGTG 540  
QY 676 GGGCAGACAATGTATTTATCTGAGGCCCTGTACCTCAAGACTCCACCTATGACGGACAT 735  
DB 541 GGGCAGACAATGTATTTATCTGAGGCCCTGTACCTCAAGACTCCACCTATGACGGACAT 600  
QY 736 ACCGTGGGCGGACTGACATGATGGGGTCTGGCCAGCTGGCAGATGGTGGTGGGGCTG 795  
DB 601 ACCGTGGGCGGACTGACATGATGGGGTCTGGCCAGCTGGCAGATGGTGGTGGGGCTG 660  
QY 796 GATGACTTTAGGAAGAGTCAGGAGTGGGCTCTGGCCAGGCTATGACTATGTGGGATGG 855  
DB 661 GATGACTTTAGGAAGAGTCAGGAGTGGGCTCTGGCCAGGCTATGACTATGTGGGATGG 720  
QY 856 AGCAACACAGCTTCTCAGTGGCTATGTGAGATGGAGTTGAGTTGACCGGCTGAGG 915  
DB 721 AGCAACACAGCTTCTCAGTGGCTATGTGAGATGGAGTTGAGTTGACCGGCTGAGG 780  
QY 916 GCCTCCAGGCTATCAGGTCACATGTAACAACATGCACAGCTGGGAGCCGCTGCCT 975  
DB 781 GCCTCCAGGCTATCAGGTCACATGTAACAACATGCACAGCTGGGAGCCGCTGCCT 840  
QY 976 GCGGGGTGGAATGTGCTTCCGGGCTGGCCCTGCGCATGGCTGGAGGGGAGCCCATG 1035  
DB 841 GCGGGGTGGAATGTGCTTCCGGGCTGGCCCTGCGCATGGCTGGAGGGGAGCCCATG 900  
QY 1036 GCGCCACACCTAGGGGGCAACCTGGGGGAGCCCGAGAGCCCGGCTGTCTAGTCCCGCTT 1095

DB 901 CGCCACAACCTAGGGGGCAACCTGGGGAGACCCAGAGCCCGGGCTGTCTAGTGGCCCTT 960  
QY 1096 GCGGGCGGTGTGGCTCGCTTTCTGCAGTGGCGCTTCTCTTTGGGGGCGCTGGTTACTC 1155  
DB 961 GCGGGCGGTGTGGCTCGCTTTCTGCAGTGGCGCTTCTCTTTGGGGGCGCTGGTTACTC 1020  
QY 1156 TTCAGCGAAATCTCTTTCATCTCTGATGTGTGAACAATCTCTCTCCGGCAGTGGGAGGC 1215  
DB 1021 TTCAGCGAAATCTCTTTCATCTCTGATGTGTGAACAATCTCTCTCCGGCAGTGGGAGGC 1080  
QY 1216 ACCTTCCCGCAGAGCCCTGTGGCGGCTGGCCACCTCCCAACACTTCAGCAGCTTG 1275  
DB 1081 ACCTTCCCGCAGAGCCCTGTGGCGGCTGGCCACCTCCCAACACTTCAGCAGCTTG 1140  
QY 1276 GAGCTGGAGCCCGAGAGGCCAGCAGCCGCTGGCCAAAGCCCGAGGGAGCCCGACCGCC 1332  
DB 1141 GAGCTGGAGCCCGAGAGGCCAGCAGCCGCTGGCCAAAGCCCGAGGGAGCCCGACCGCC 1197

RESULT 4  
US-08-456-647B-19  
; Sequence 19, Application US/08456647B  
; Patent No. 5811516  
; GENERAL INFORMATION:  
; APPLICANT: Lemke Ph.D. et al., Greg E.  
; TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: US  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456.647B  
; FILING DATE: 02-JUN-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/237,401  
; FILING DATE: 02-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/884,486  
; FILING DATE: 15-MAY-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wetherell Ph.D., John R.  
; REGISTRATION NUMBER: 31,678  
; REFERENCE/DOCKET NUMBER: 07251/007002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 678-5070  
; TELEFAX: (619) 678-5099  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3120 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
; IMMEDIATE SOURCE:  
; CLONE: Tyro-10  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 485..3047  
; US-08-456-647B-19

Query Match

18.4%; Score 669.4; DB 3; Length 3120;

Best Local Similarity 56.98: Pred No 1 4e-139.

2000 Local Similarity 50.5%, Freq.: NO. 1.4e-139;  
 Matches 1553; Conservative 0; Mismatches 986; Indels 192; Gaps 10;

QY	109	CTGCTGCTGCTCTTTGGTGGCAAGTGGAGATGCTGACATGAAGGACATATTTTATCCTGCC	168
DB	110		
DB	509	CTGTGCTGCTCTGCTCTTGTCTAFCCTGGGTTCTGCAAAAGCTCAGGTTAATCCAGCC	568
QY	169	AAGTGGCGCTATCCCTGGSCATGCAGGACCGACCATCCAGACAGTGCATCTCTGCT	228
DB	170		
DB	569	ATATGCCGCTATCCTCTGGGATGTCAGGAGGCCACATTCAGATGAGGACATCACGCC	628
QY	229	TCCAGTCTCTGTGCAGATTCCACTGCGGCCGCCACAGCAGCTTTGGAGAGCAGTGACGGG	288
DB	230		
DB	629	TCAAGTCAGTGGTCAGAATCCACGGCTGCCAAATATGGGAGGCTGCATCTCGAAGAAG	688
QY	289	GATGGGGCCTGTGCCCGCAGGGTCGGTGTTCCTCCAAAGGA---GGAGGAGTACTTGCAG	345
DB	290		
DB	689	GATGGAGCCTGTGTCTCTGAGATTCCAGTTCAGTCAACCCGATGACTCGAAGGAAATTTCTGCAG	748
QY	346	GTGGATCTACAAGCACTGCACCTGTGGCTCTGTGGGACCCAGGAGCGCATGCCGGG	405
DB	347		
DB	749	ATTGACTTGGAACTTACACTTTATCACTCTTGTGGGAGCCAGGCGGCCATGCAGGG	808
QY	406	GGCCTGGGCAAGGATTTCTCCGGAGCTACCGGCTGCGTTACTTCCCGGAGTGGTCGCGC	465
DB	407		
DB	809	GGTCATGGCATTAATTTGACCCCATGTTACAAGATCAACTACAGTCGGATGGCAGTCGC	868
QY	466	TGATGGGCTGGAGGACCGCTGGGCTCAGGAGGTGATCTCAGGCAATGAGACCTTGAG	525
DB	467		
DB	869	TGATCTCTGGCGTAAACCGCATGGGAAGCAGTGCTTGATGGAACAGTAACCCCTAT	928
QY	526	GGAGTGTGCTGAAGCACTTGGGCCCCCATGTGTGCCGACTGGTTCGGTCTCTACCCC	585
DB	527		
DB	929	GATGATTTCTGAAGGACTTGGAGCCACCATCTGCGCAGATTTGTTCGGCTTATCCCA	988
QY	586	CGGGCTGACCGGCTCATGAGCGTCTGTCTCGGGGTAGAGCTCTATGGCTGCTCTGGAGG	645
DB	587		
DB	989	GTCACTGACCACTCCATGAACGCTGTCATGAGGTTGAGCTTTATGTTGTCTGGGCTA	1048
QY	646	GATGGACTCTGTCTTACACCGCCCTCTGGGGCAGACAAATGATTATCTGAGG-----	701
DB	647		
DB	1049	GATGGCTTGGTATCCTTACAATGCTCCAGCTGGACAGCAAGTTGTACTCCTGGAGGCTCC	1108
QY	701	-CGGTACTCAACGACTCCACCTATGACGACATACCGTGGGCGGACTGCAGTATGSGG	759
DB	702		
DB	1109	ATCATTTATCTGAATCTGCTATGATGAGCTGTTGGTACAGCATGACTGAGGG	1168
QY	760	GGTCTGGGCAAGTGCAGATGGTGTGGGGCTGGATGATCTTAGGAAGAGTCAAGAG	819
DB	761		
DB	1169	---CTAGGCGAGTTGACTGATGAGGATATCCGGCCTGGATGATTTTACCAGACCCATGA	1225
QY	820	CTCGGGTCTGGCAGGCTATGACTATGGGATGGAGCAACCAACAGCTTCTCCAGTGGC	879
DB	821		
DB	1226	TACCAGCTGGGCTGGCTATGACTACGTGGGATGGGGAAGAAAGTGTACCAACGGT	1285
QY	880	TATGTGAGATGGAGTTTGAGTTTGACCGGCTGAGGGGCTTCCAGGCTATGAGGTCCAC	939
DB	881		
DB	1286	TTCAATTGAGATCATGTTTGAATTTGACCGAATCAGGAATTTTACTTACCATGAAGTCCAC	1345
QY	940	TGTAAACAATGCACAGCTGGGAGCCCTGCTGCTGCGGGGTGGAATGTCGCTTCCGG	999
DB	941		
DB	1346	TGCAACAACATGTTTCTTAAGGTGTGAAGATTTTAAAGAGGTCAGCTGCTACTTTCGC	1405
QY	1000	CGTGGGCCCTGCCATGCCCTGGAGGGGAGCCCATGCCACAACTAGGGGGCAACCTG	1059
DB	1001		
DB	1406	TCGG---AGCCAGCAGTGGGAACCCATGCTGTCTACTTTCCCTGGTCTGGAGCAT	1462
QY	1060	GGGAGCCCCAGAGCCCGGCTGTCTCAGTGCCTTGGCGGCGGTGTGGCTTCTG	1119
DB	1061		
DB	1463	GTGAACCCAGTGCCTGGTTGTACGGTGGCCCTCCACCACCGAATGCCAGTGCCTATC	1522
QY	1120	CAGTGGCGCTTCTCTTTTGGGGGCGCTGGTTACTTCTTACGGGAAATCTCCTTCTCT	1179
DB	1121		

[illegible]

QY 2257 GCGCTGGGACGGGAGGCTGGCGAGGGGCCACCATCAGCTACCCAAATGCTGTCAT 2316  
Db 2502 -----GTTCTCTTCTAGTGATGCCACAGTCAGTTACGCCAACCTGAAGTTT 2548  
QY 2317 GTGGCAGCCAGATCGCTCCGCGATCGGTATCTGGCCACACTCACTTTGTACATCG 2376  
Db 2549 ATGGCAACCCAGATGCTGCTGATGAGTACCTTTGCTCTCACTTTGTCCACCGA 2608  
QY 2377 GACTGGCCAGCGGAACTCCCTAGTTGGGAAATTTCAACATCAAAATCGCAGACTTT 2436  
Db 2609 GATCTGGCCACAGAACTCTTTAGTGGCAAGAAATACCAATCAAGTATGCTGATTT 2668  
QY 2437 GGCATGAGCCGGAACCTCTATGCTGGGACTATTACCGTGTGCGAGGGCCGCGAGTGTG 2496  
Db 2669 GGCATGAGCAAAACCTGTACAGTGTGATTTACTACCGGATCCAGGGCCGGGGTGTCTC 2728  
QY 2497 CCATCCGCTGGATGGCTGGGAGTGCATCTCATGGGAAATTCACGACTCGGAGTGAC 2556  
Db 2729 CCCATTCGCTGGATGCTCGGGAAGCATCTTCTGGCAAAATTCACCAAGCAAGTGTAT 2788  
QY 2557 GTGGGCGCTTTGGTGTGACCTGTGGGAGGTGCTGATGCTCTGTAGGGCCCGCCCTTT 2616  
Db 2789 GTGGGCGCTTTGGGAGTCTGTGGGAGACCTTCACCTTTTGGCAGGAGCCCTAT 2848  
QY 2617 GGGCAGCTACCGCAGCAGCAGGTATCTGAGAACGCGGGGAGTCTTCCGGGACCAAGGCG 2676  
Db 2849 TCCAGCTGTGGATGAGCAGGTATTCGAGAACACTGGAGAGTCTTCCGAGACCAAGGG 2908  
QY 2677 CGGAGCTGTACCTGTCCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2736  
Db 2909 AGCAGATCTATCTCCCTCAACAGCCCTTTGGCCGAGCTCTGTGTATAGCTGATGCTC 2968  
QY 2737 CGGTGCTGGAGCGGGAGTCTGAGCAGCAGCAGCCCTTTTCCAGCTGCTGCTGCTGCTG 2796  
Db 2969 AGCTGCTGGAGAGAGAACCAAGCAGCCGCGCATCTTCCAGGAATACACCTCTGCTT 3028  
QY 2797 GCAGAGATGACTCAACACGGTGTGAATCA 2827  
Db 3029 CTTCAAGAGGCGGAGTGTATGATGATCA 3059

## RESULT 5

US-08-237-401A-19  
; Sequence 19, Application US/08237401A  
; Patent No. 5837448  
; GENERAL INFORMATION:  
; APPLICANT: Lemke Ph.D. et al., Greg E.  
; TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: US  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/237,401A  
; FILING DATE: 02-MAY-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/884,486  
; FILING DATE: 15-MAY-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Haile Ph.D., Lisa A.  
; REGISTRATION NUMBER: 38,347  
; REFERENCE/DOCKET NUMBER: 07251/007001

## TELECOMMUNICATION INFORMATION:

; TELEPHONE: (619) 678-5070  
; TELEFAX: (619) 678-5099  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3120 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
; IMMEDIATE SOURCE:  
; CLONE: Tyro-10  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 485..3047  
; US-08-237-401A-19

Query Match 18.4%; Score 669.4; DB 3; Length 3120;  
Best Local Similarity 56.9%; Pred. No. 1.4e-139;  
Matches 1553; Conservative 0; Mismatches 986; Indels 192; Gaps 10;

QY 109 CTGCTGCTGCTCTTGGTGGCAAGTGGAGATGCTGACATGAAGGACATTTTGTATCTCTGCC 168  
Db 509 CTGCTGCTGCT 568  
QY 169 AAGTGGCCCTATGCGCTGGGATGACAGGACCGGACCATCCAGACAGTACATCTCTCTCTCT 228  
Db 569 ATATGGCCCTATCTCTCTGGGATGTCAGGAGGCGCCACATTCAGATGAGGACATCACAGCC 628  
QY 229 TCCAGCTCTCTGTCAGATTCACATGCGCGCGCCGCGCCACAGCAGCTGGAGAGCAGTACGGG 288  
Db 629 TCAAGTCACTAGTGGTCAAGATCCAGGCTGCCAAATATGGAGGCTGGACTCTGAAGAAGGA 688  
QY 289 GATGGGCTGCTGCTGCGCCGAGGCTCGGTGTTTCCCAAGGA---GGAGGAGTACTTTCAG 345  
Db 689 GATGGAGCTGCTGCT 748  
QY 346 GTGGATCTTAAACAGCTGACCTCTGCTGGCTCTGCTGGGACCCAGGAGCGCATGCCCGG 405  
Db 749 ATTGACTTTCGAACCTTACACTTATCACTCTTGTGGGAGCCAGGCGCGCATGTCAGAGG 808  
QY 406 GGCCTGGGCAAGGAGTCTCTCCGAGCTACCGGCTGGCTTACTCCCGGAGTGTCTGCGCCG 465  
Db 809 GGTCAATGCACTTGAATTTGCACCATGTACAAGATCAACTACAGTCGGGATGGCAGTCGC 868  
QY 466 TGGATGGCTGGAAGGAGCCGCTGGGTCAGGAGGTGATCTCAGGCAATGAGGACCTTGA 525  
Db 869 TGGATCTCTGCTGCTTACCGGATGGGAGCAGGTGCTTGTATGGAACAGTAAACCTTAT 928  
QY 526 GGAGTGTGCTGAAGGACCTTTGGGCCCCCCTATGCTGGCCGACTGGTTCGCTTCTACCCC 585  
Db 929 GATGTATTCTCTGAAGGACTTGGAGCCACCATCTGCGCAGATTTGTTCGCTTATCCCA 988  
QY 586 CGGCTGACCGGGTCACTGAGGCTGCTGCTGGGCTAGAGCTCTATGCTGCTCTGAGG 645  
Db 989 GTCAGTGACACTCCATCAACGCTGTGATGAGGTTTATGCTTGTGCTGCTGCTA 1048  
QY 646 GATGACTCTCTCTTACACCCGCTTGGGCGAGCAATGATTTATCTGAGG-----701  
Db 1049 GATGCTTGGTATCCTACAATGCTCCAGCTGGACAGCAGTTTGTACTCCCTGGAGGCTCC 1108  
QY 701 -CCGTGTACCTCAAGACTCCACCTATGACGAGACATACCGTGGGCGGAGTCACTATGGG 759  
Db 1109 ATCAATTATCTGAATGATTTCTCTATGATGAGCTGTGGGTACAGCATGACTGACTGAAGG 1168  
QY 760 GGTCTGGGCGGAGCTGGCAGATGCTGTGGGCTGGATGACTTTAGGAAGAGTCAGAG 819  
Db 1169 ---CTAGGCGAGTTGACTGATGGAGTATCCGCGCTGGATGATTTTACCAGACCCATGAA 1225  
QY 820 CTGCGGCTCTGGCCAGGCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 879  
Db 1226 TACCACGTGTGCGCTGGCTATGACTACGTACGTGCGGATGGCGGAAGAAAGTGTCTACCAACGT 1285

QY 880 TATGTGAGATGAGTTTTCAGTTTGGAGGCGCTGAGGGCCCTCCAGGCTATGAGTGCAC 939  
DB 1286 TTCAATGAGATCATGTTTGAATTTGACCGAATCAGGAATTTTACTACCATGAGGATCCAC 1345  
QY 940 TGTAAACAATGACACACGCTGGAGCCGCTCTGCTGCGGGGGTGAATGTCGCTTCCGG 999  
DB 1346 TGCACAACATGTTTCTAAAGGTGTGAAGATTTTAAAGAGGTCCAGTCTACTTTCGC 1405  
QY 1000 CGTGGCCCTGCGCATGCGCTGGAGGGGGAGCCCATGCGCCACAACCTAGGGGCAACCTG 1059  
DB 1406 TCGG---AAGCCAGCGAGTGGGAACCACTGCTGTCTACTTTCCCTCTGCTCGACCAT 1462  
QY 1060 GGGGACCCAGAGCCGCGCTCTCTCAGTGCCTTGGGGCCGCTGGCTGCTGCTTCTG 1119  
DB 1463 GTGAACCCAGTGGCCGCTTGTACGCTGGGCTTCCACCAACCAATGGCCAGTGCCTATC 1522  
QY 1120 CAGTGGCCGCTTCTCTTTCGGGGGCGCTGTTACTTCTCAGGGAATCTCTTCTCATCTCT 1179  
DB 1523 AAGTGCATACCATTTTCCGACAGGTGGATGATGTTTCAGGAGATCACTTTCCAAATCA 1582  
QY 1180 GATGGGTGAACAATTCCTCTCCGGACTGGAGGACCTTCCCGCCAGCCCTGCTGCTG 1239  
DB 1583 GATGCTGCAATGTATAACAATCTGGAGCCCTTCCACCTCTCTTA----- 1629  
QY 1240 CCGCTGGCCCACTCCCACTTTCAGCACTTGGAGCTGGAGCCCAAGAGCCAGGAG 1299  
DB 1629 -----TGCCACCCACCACTATGAT 1648  
QY 1300 CCGTGGCAAGCCGAGGGAGCCGACCGCATCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1359  
DB 1649 CCATGCTTAAAGTTGATGATAGCAACTCGGATCTCTGATGTTGTTGGTGGCCATC 1708  
QY 1360 ATCTGCTCTGCTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1419  
DB 1709 ATCTGCTGCTGCTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1768  
QY 1420 CTCAGCAAGCTGAAGGAGGTGTTGAGAGGAGCTGACGGTTTCACTCTCTGCTGCT 1479  
DB 1769 CTAGAAAGCTTTCAGGAGGATGCTGATGATGAATGACAGTCAAGCTTCTGCTGCTGCT 1828  
QY 1480 GGGGACACTATCTCATCAACACCGCCAGGTCTCTAGAGAGCCACCCCTGACAGGAG 1539  
DB 1829 AGCGAGTCCAGCATGTTCAATAACACCGCTCTC-----ATCACCAAGTGAACAGGAG 1882  
QY 1540 CCGGCGCTCTGGGAATCCGCCCCACCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1599  
DB 1883 TCCAATCTACTTATGATCGAATCTTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1942  
QY 1600 CTGCTCTCAATCCAGCTTACCGCTCTCTCTGCGCACTTACGCGCTGCTGCTGCTGCT 1659  
DB 1943 CTGATC----- 1949  
QY 1660 CCGGCGCCCCCACACCGCCCTGGGCAACACCAACACCCAGGCTTACAGTGGGAG 1719  
DB 1949 -----CGAAGCTTCCAGTTTGTCCAGGAGGAGGA 1983  
QY 1720 TATATGGAGCTGGAAGCCAGGCGCCCGCTTGTGCCCCACCTTCCCGCAACAGCGTCT 1779  
DB 1984 GTCAAGGTGCTGCTGTTGTTGAAGCCGCT-----CAGCCCAATGGACCTGAGGGGCTG 2038  
QY 1780 CCCCATTATCGGAGCTGACATTTTACCTGAGGCGCTCACCGGGGGCAACACCTAT 1839  
DB 2039 CCCCATTATGAGAGCGGACATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2098  
QY 1840 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1896  
DB 2099 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2158  
QY 1897 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1956  
DB 2159 CCGAGGAACCTGTTGGCTTCAAGGAGAGCTGGGAGAGCCAGTGTGGGAGGTTTCT 2218

QY 1957 CTGTGTGAGTTCGACAGCCCTCAAGATCTGGTCACTTCTGATTTCCCTTAAATGTGCTG 2016  
DB 2219 CTCTGTGAAGTGGAGGAATGGAATAATCAAGACAAAGATTTGCACCTAGATGTCAGT 2278  
QY 2017 AAGGGACACCTTTCGT 2076  
DB 2279 GCCAACCAGCTCTCTCTGTTGGCGGTGAAATGCTCCGAGCAGATGCCAACAAGATGCC 2338  
QY 2077 AGGAATGATTTCTTGAAGAGGTGAAGATCATGTGAGGCTCAAGGACCAACACATCAT 2136  
DB 2339 AGGAATGATTTCTTGAAGAGATCAAGATCATGTCTGGCTCAAGGACCAACACATCAT 2398  
QY 2137 CGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2196  
DB 2399 CGTCTCTTAGT 2458  
QY 2197 AAGGGACACCTCAACCACTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2256  
DB 2459 AATGGAGATCTTAAATCAGTTTCTTCTGCGCACGAGCTCTGA----- 2502  
QY 2257 GCGCTGGGACCGGAGGCTGCGCAGGCGCCACCATCAGCTACCAATGCTGCTGCAT 2316  
DB 2502 -----GTTCCCTGTTCTAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2548  
QY 2317 GTGGACGCCAGATGCGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2376  
DB 2549 ATGCAACCCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2608  
QY 2377 GACTGGCCAGCGGAACTCTATGCTGGGAGTATTTAGTGGGAAATTTACCAATCAAAATCGCAGACT 2436  
DB 2609 GATCTGGCCACAGAACTGTTTGTGGGCAAGAAATTTACCAATCAAAATCGCAGACT 2668  
QY 2437 GGCATGAGCGGAACTCTATGCTGGGAGTATTTAGTGGGAAATTTACCAATCAAAATCGCAGACT 2496  
DB 2669 GGCATGAGCAAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2728  
QY 2497 CCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2556  
DB 2729 CCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2788  
QY 2557 GTGTGGCCCTTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2616  
DB 2789 GTGTGGCCCTTTCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2848  
QY 2617 GGCAGCTCACCCAGCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2676  
DB 2849 TCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2908  
QY 2677 CGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2736  
DB 2909 AGCAGATCTATCTCCCTCAACAGCCCTTTCGCGGACTGCTGCTGCTGCTGCTGCTGCTGCT 2968  
QY 2737 CGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2796  
DB 2969 AGCTGCTGGAAGAGAGAAACCAAGCACCAGCCATCTCTCCAGGAAATACACCTCTGCT 3028  
QY 2797 GCAGAGATGCTCACTCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2827  
DB 3029 CTTCAAGGAGGAGGAGTGTGATGCTCA 3059

## RESULT 6

US-08-336-343A-3

; Sequence 3, Application US/08336343A

; Patent No. 5677144

; GENERAL INFORMATION:

; APPLICANT: Ullrich, Axel

; APPLICANT: Alves, Frauke

; TITLE OF INVENTION: CCK-2, A No. 5677144el Receptor Tyrosine Kinase

; NUMBER OF SEQUENCES: 43

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie &amp; Edmonds

; STREET: 1155 Avenue of the Americas

CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/336,343A  
FILING DATE: 08-NOV-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7683-065  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3157 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: unknown  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 370..2934  
US-08-336-343A-3

Query Match 18.38; Score 566.8; DB 2; Length 3157;  
Best Local Similarity 56.88; Pred. No. 5.3e-139;  
Matches 1532; Conservative 0; Mismatches 977; Indels 189; Gaps 9;

QY 110 TGCTCTGCTCTTGGTGGCAAGTGAGATGCTGACATGAAGGGACATTTTGATCCTGCCA 169  
DB 395 TGGTCTGCTCTGCTGCTGCTATCTTGAGTTCTGCAAAAGCTCAGTTTAATCCACTA 454  
QY 170 AGTGGCGGTATGCCCTGGGATGACGAGACCGGACCATCCAGACAGTGCATCTCTCTT 229  
DB 455 TATGCCGCTATCTCTGGGCTATGTCAGGAGCCAGATTCCAGATGAGGATCATCAGCTT 514  
QY 230 CCAGCTCTGCTCAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 289  
DB 515 CCAGTCACTGGTCAAGTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 574  
QY 290 ATGGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 346  
DB 575 ATGGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 634  
QY 347 TGGATCTTACAAGACTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 406  
DB 635 TTGACTTGCACACCTCCATTTTATCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 694  
QY 407 GCTGGCAAGAGTTCTCCCGAGCTACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 466  
DB 695 GTGATGGATCAGTTTGGCCCAATGTAAGATCAATTAAGTGGGATGGGACTCGCT 754  
QY 467 GGATGGGTGGAAGGACCGCTGGGCTCAGGAGGTGATCTCAGGCAATGAGGACCTGAGG 526  
DB 755 GATCTCTGGGGAACCGCTCATGGGACAGGCTGCTGATGGAATAGTAACCCCTATG 814  
QY 527 GAGTGGTCTGAAGGACCTTGGGCCCCCATGTTGCCCGACTGCTGCTTCTACCCCC 586  
DB 815 ACAATTTCTTAAAGGACTTGGAGCGCCCATTTAGCCAGATTTTGTCCGTTCTATCCAG 874  
QY 587 GGGCTGACCGGGTCAATGAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 646

DB 875 TCACCGACCACTCCATGAATGTGTGATGAGATGGAGCTTTACGGTGTCTGTGGCTAG 934  
QY 647 ATGGACTCTCTGTACACCGCCCTGTGGGCGACAGAAATGATTTATCTGAGG----- 701  
DB 935 ATGGCTTGGTGTCTTACAAATGCTCCAGCTGGGACGAGTTTGTACTCCCTGGAGTTCCA 994  
QY 701 CGGTGTACTCTAACGACTCCACCTATGACGACATACCGTGGGCGGACTGCAAGTATGGG 760  
DB 995 TCATTTATCTGAATGATTTCTGTATGATGAGCTGTGGATGATGAGTGTGGACAGATGAC 1054  
QY 761 GTCTGGCCAGCTGGCAGATGCTGTGGGCTGGGACTTTAGGAAGAGTGCAGAGC 820  
DB 1054 --CTAGGCCAATTGACCGATGTGTGCTGCTGGCTGGACGATTTTACCAGACCCATGAAT 1111  
QY 821 TCGGGTCTGGCCAGGCTATGACTATGTGGATGGAGCAACACAGCTTTCTCCAGTGGCT 880  
DB 1112 ACCACGTGTGGCCCGCTATGACTATGTGGCTGGCGAAGAGAGTGCACCAATGGCT 1171  
QY 881 ATGTGAGATGGAGTTTGAATTTGACCGCTGAGGGCTTCCAGGCTATGCAAGTCCACT 940  
DB 1172 ACATTTGAGATCATGTTTGAATTTGACCGCATCAGGAATTTCACTACCATGAAGTCCACT 1231  
QY 941 GTACAACATGCACACGCTGGGAGCCGCTCTGCTGGGCGGTGGAATGCTCGCTTCCGCG 1000  
DB 1232 GCAACAACATGTTTCTAAAGTGTGAAGATCTTTAAGAGGTACAGTGTCTACTTCCGCT 1291  
QY 1001 GTGGCCCTGCCATGGCTTGGAGGGGAGCCCATGCGCCCAACAACTAGGGGCAACCTGG 1060  
DB 1292 CTG---AAGCCAGTGAAGTGGGAACCTAATGCAATTTCTTCCCTTCTCTGGATGAG 1348  
QY 1061 GGGACCCAGAGCCCGGCTGCTCAGTGCCTTGGGCGGCTGTGGCTGCTGCTTCTGCTG 1120  
DB 1349 TCAACCCAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1408  
QY 1121 AGTGGCGCTTCTCTTGGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1180  
DB 1409 AGTGTCAATACCATTTGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1468  
QY 1181 ATGTGGTGAACAATTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1240  
DB 1469 ATGTGCAATGTACAACAACCTCTGAAGCCCTGCCACCTCTCTA----- 1514  
QY 1241 CGCTGGCCCACTCCACCACTTCCAGAGCTTGGAGCTTGGAGCTTGGAGCTTGGAGCTT 1300  
DB 1514 -----TGGCACCCACACCTATGATC 1534  
QY 1301 CGGTGGCCAAAGCCCGAGGGAGCCGAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1360  
DB 1535 CAATGCTTAAAGTTGATGACAGCAACACTCGGATCTGATTTGGCTGCTTGGTGGCCATCA 1594  
QY 1361 TCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1420  
DB 1595 TCTTTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1654  
QY 1421 TCAGCAAGCTGACGGAGGGTGTGGAAGAGAGCTGACGCTTCACTCTCTGCTGCTGCTG 1480  
DB 1655 TGGAGAGGCTTCTCGGAGGATGCTGAGATGATGAATGACAGTCAAGCTTTTCCCTGCA 1714  
QY 1481 GGGACACTATCTCATCAACAACCGCCAGGTCCTAGAGAGCCACCCGCTTACCAAGAGC 1540  
DB 1715 GTGATTTAGCATGTTCACAAT-----RACCCCTCT 1747  
QY 1541 CCGGGCTCGTGGGAATCCGCCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1600  
DB 1748 CATCACTAGTGAACAAGGGTCCAACTGACTTACGATGCTGCTGCTGCTGCTGCTGCTGCT 1795  
QY 1601 TGCTCTCCAATCCAGCTACCGCTTCCCTTCTGCGCACCTTACCGCCGCTCCCTCGAGGC 1660  
DB 1795 -----CCCCCTCGGC 1804  
QY 1661 CGGGCCCCCACCACCGCTGCGGCAACACCCAGGCTTACAGTGGGGACT 1720

Db 1805 CTGACTACAGGAGCCATCCAGGCTCATAGGAAACTCCAGAAATTTGCTCCAGGGAGG 1864  
QY 1721 ATATGAGCCTGAGAAGCCAGGCGCCCGCTTCTGTCGCCACCTCCCCAGAAACAGCGTCC 1780  
Db 1865 AGGAGTCAAGCTCGAGGGTGTGTGAAGCCAGTCCAGCCAGTGGCCCTGAGGGGTGC 1924  
QY 1781 CCATATGCGGAGGCTGACATTTTACCTCGAGGGGTACCGGGGGCAACACCTATG 1840  
Db 1925 CCACATGAGAGGCTGACATAGTAAGCTCCAGGAGTGCAGGAGGCAACACATACT 1984  
QY 1841 CTGTGCTGCACTGCCCCCAGGGGAGTCC---GGGATGGGCCCCCAGAGTGGATTTCC 1897  
Db 1985 CAGTGCCTGCCCTCACCATTGGACCTCTCTCAGGAAAGATGTGGTGGAGGATTTCC 2044  
QY 1898 CTCGATCTCGACTCCCTTCAAGGAGAAGCTTGGAGGGCCAGTTTGGGGAGGTGCACC 1957  
Db 2045 CCAGGAAACTCTCACTTTTCAAGGAGAAGCTGGGAGAAGGAGTTTGGGGAGGTTCATC 2104  
QY 1958 TGTGTCAGTCCGACAGCCCTCAGATCTGGTCAGTCTTGAATTTCCCTTATGTGGTA 2017  
Db 2105 TCTGTGAAGTGGAGGAATGGAAATTCAGAGCAAAAGATTTGCCCTAGATGTCACTG 2164  
QY 2018 AGGACACCCCTTGTCTGTAGTGTCAAGATCTTACGGCCAGATGCCACCAAGATGCCA 2077  
Db 2165 CCAACAGCCTCTCTGTGTGGTGTGAATAATGCTCGGACGATGCCAACAAAGATGCCA 2224  
QY 2078 GGAATGATTTCTTGAAGAGGTGAAGATCATGTGAGGCTCAAGGACCAACACATCATTC 2137  
Db 2225 GGAATGATTTCTTGAAGAGATGAAGATCATGTCTCGGCTCAAGGACCAACACATCATTC 2284  
QY 2138 GGTGCTGGGCTGTGTGTGTCAGGAGGACCCCTCTGCATGATTTACTGACTACATGGAGA 2197  
Db 2285 ATCTATTATCTGTGTATCACTGTAGTACCCCTCTGTATGATCACTGAATACATGGAGA 2344  
QY 2198 AGGCGACCTCAACCACTTCTCAGTGCCACCAGCTGAGGACAGCAAGCGCCGAGGGG 2257  
Db 2345 ATGGAGATCTCAATCAGTTCTTCCCGCACAGGCCCTTAATTTCTCTCCAGCG--- 2402  
QY 2258 CCCTGGGACGGGAGGCTGCGAGGGGCCACCATCAGCTACCAATGCTGTGTCATG 2317  
Db 2402 -----ATGTACGCACTGTCTGATTTACACCAATCTGAAGTTTA 2437  
QY 2318 TGCGAGCCAGATCGCTCCGCGATCGCGTATCTGGCCACACTCAACTTTGTACATCGGG 2377  
Db 2438 TGCTACCCAAATGGCTCTGCGATGAAGTACCTTTCTCTTAATTTTGTTCACCGAG 2497  
QY 2378 ACTGCGCCACGGGAACCTGCTAGTTGGGAAATTTTACCACATCAAAATCGAGACTTTG 2437  
Db 2498 ATCTGCGCCACAGAACTGTTAGTGGGTAGAACTACAAATCAAGATAGCTGACTTTG 2557  
QY 2438 GCATGAGCGGAACCTCTATGCTGGGACTATTACCGTGTGCGAGGGCGGCACTGTCTGC 2497  
Db 2558 GAATGAGCAGGAACCTGTACAGTGTGACTATTACCGGATCCAGGGCGGCGAGTGTCTC 2617  
QY 2498 CATCGCTGGATGCTGGGAGTGCATCTTCATCGGGAAGTTACGACTCGGAGTACG 2557  
Db 2618 CTATCGCTGGATGCTGGGAGATGATCTTGTCTGGGCAAGTTTCACTACAGCAAGTATG 2677  
QY 2558 TGTGGGCTTTTGTGTGACCTCTGGGAGGTGCTGATGCTCTGTAGGGCCCGAGCCCTTTG 2617  
Db 2678 TGTGGGCTTTTGGGTTACTTGTGGGAGACTTTACCTTTTGTCAAGAACAGCCCTATT 2737  
QY 2618 GCGAGCTACCGCAGCAGGAGTTCATCGAAGCGGGGAGTTCTTCGGGACCAAGGCGCC 2677  
Db 2738 CCAGCTGTGATGAACAGGTTATTGAGAACTAGTGGAGAGTTCTTCGAGACCAAGGGA 2797  
QY 2678 GCGAGCTGTACCTGTCCCGGCGCTCTGCCCGCAGGGCTATATGAGTGTATGCTTC 2737  
Db 2798 GCGAGCTGTACCTGTCCCGCAGGCAATTTGTCTGACTCTGTGTATAGCTGATGCTCA 2857  
QY 2738 GGTGTGTCAGCGGAGTCTGAGCAGGACCAACCTTTTCCAGCTGATCGTCTCT 2795  
Db 2858 GCTGCTGGAGAGATAGGAGAACCGTCTCTCAATTCAGAAATCCACCTTCTGCT 2915

## RESULT 7

US-08-336-343A-5/C  
; Sequence 5, Application US/08336343A  
; Patent No. 5677144  
; GENERAL INFORMATION:  
; APPLICANT: Ullrich, Axel  
; APPLICANT: Alves, Frauke  
; TITLE OF INVENTION: CK-2, A No. 5677144el Receptor Tyrosine Kinase  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/336,343A  
; FILING DATE: 08-NOV-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7683-065  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-9741/8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3157 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: unknown  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
US-08-336-343A-5

Query Match 18.3%; Score 666.8; DB 2; Length 3157;  
Best Local Similarity 56.8%; Pred. No. 5.3e-139;  
Matches 1532; Conservative 0; Mismatches 977; Indels 189; Gaps 9;

QY 110 TGCTGCTGCTCTTGGTGGCAAGTGGAGATGCTGACATGAAGGAGCAATTTTATCTCTGCCA 169  
Db 2763 TGTGCTGTTTCTTGGCTGCTGCTTATCTTGAATTTCTCAAAAGCTCAGGTTAATCTCAGCTA 2704  
QY 170 AGTCCGCTATGCTGCTGGGATCGAGGACCGGACCATCCAGACAGTGCATCTCTGCTT 229  
Db 2703 TATGCGCTATCTCTTGGCATGTCAGGAGGCCAGATTCAGATGAGGACATCAGACTT 2644  
QY 230 CAGCTCTCTGCTCAGATTCACATGCGCCGCCACAGCAGGTTGGAGAGCAGTGCAGCGGG 289  
Db 2643 CCACTCAGTGTGTCAGAGTCCACAGCTGCCAATATGAAAGGCTGGACTCAGAAGAAGGGG 2584  
QY 290 ATGGGCGCTGCTGCGCCGAGGCTCGGTGTTCCTCCCAAGGA---GGAGGAGTACTTTCAGG 346  
Db 2583 ATGAGCGCTGGTGGCTGAGATTCAGTGGAACTGATGATGAGGAGTTTCTGCAGA 2524  
QY 347 TGGATCTACAACACTGCACTGCTGGTGGCTCTGTGGGACCCAGGAGCGCATGCGGGG 406  
Db 2523 TTGACTTGACACCCCTTCAATTTATCATTCTGTGGGGACCCAGGCGGCCCATGAGGAG 2464  
QY 407 GCTTGGGCAAGGAGTTCTCCCGAGGCTACCGGCTGCTTACTCCCGGAGTGGTCCCGCT 466



Db 2463 GTCATGGCATCGAGTTGGCCCGCCATGTACAAAGATCAATTTACATCGGGATGGCACTCGCT 2404  
Qy 467 GCGTGGGCTGGAAGAACCGCTGGGGTCAAGAGGTGATCTCAGGCAATGAGAACCCCTGAGG 526  
Db 2403 GGAATCTCTGGCGGAACCGTCATCGGAACACAGGTGCTGGATGGAATAGTAAACCCCTATG 2344  
Qy 527 GAGTGGTCTGAAGGACCTTGGGCGCCCATAGTGTGGCCGACATGGTTCCGCTTCTACCCCC 586  
Db 2343 ACAATTTCTTAAGAGACTTGGAGCGGCCCATTTAGCCAGATTTGTCGGGTTCATTCCAG 2284  
Qy 587 GGGTGCACGGGTGATGAGCGTCTGTCTGGGTAGAGCTCTATGGCTGCTCTGAGGG 646  
Db 2283 TCACGCACTTCCATGAATGTGTATGATGAGATGGAGCTTTACGGCTGTCTGGCTAG 2224  
Qy 647 ATGACTCTCTTACACCCCGCTGTGGGCGACAAATGATTTATCTGAGG----- 701  
Db 2223 ATGGCTTGGTCTTACAACTCTCCAGCTGGCGAGCAGTTTGTACTCCCTGGAGTTCCA 2164  
Qy 701 CCGTGTACTCAAGACTTCCACCTATGACGACATACCGTGGGCGGACTCGAGTATGGG 760  
Db 2163 TCAATTTCTGAATGATTTCTGTCTATGATGAGCTGTGGATACAGCATGACAGAAGG- 2104  
Qy 761 GTCTGGCGGCTGCGAGATGCTGTGGGTGGCTGGATGACTTTAGGAAGAGTCAGGAGC 820  
Db 2104 -CTAGGCGCAATGACCCATGCTGATGCTGGCTGGAGATTTACCCAGACCCATGAAT 2047  
Qy 821 TCGGGTCTGGCGGCTATGACTATGTGGATGGAGCAACACAGCTTCTCCAGTGGCT 880  
Db 2046 ACCAGTGTGGCGGCTATGACTATGTGGCTGGCGAAGAGTGGCCACCAATGGCT 1987  
Qy 881 ATGTGGAGATGGATTTGAGTTTGACCGGCTGAGGGCTTCCAGGCTATGAGGTCCACT 940  
Db 1986 ACAATGAGATCATGTTTGAATTTGACCCCATCAGGAATTTCACTACCATGAAGTCCACT 1927  
Qy 941 GTAAACAATCACACGCTGGGAGCCGCTGTCTGGCGGGTGGAAATGCTGCTTCGGG 1000  
Db 1926 GCAACAACATGTTTCTAAAGGTGGAAGATCTTTAAGGAGGTACAGTCTACTTCGCT 1867  
Qy 1001 GTGGCCCTGCCATGGCTGGAGGGGAGCCCATCGCCACAACTTAGGGGGCAACCTGG 1060  
Db 1866 CTG---AAGCCAGTGAAGTGGGAACCTAATGCCATTTCTTCCCGCTTGTCTGGATGACG 1810  
Qy 1061 GGGACCCAGAGCCGGCTGTCTCAGTGGCCCTTGGCGGCTGTGGCTGCTTTCTGC 1120  
Db 1809 TCAACCCAGTGTCTGGTTTGTACGGTGGCTCTCCACACCGAATGGCCAGTGGCCATCA 1750  
Qy 1121 AGTGGCTTCTCTTTGGGGGCTGGTACTCTTTCAGCGAAATCTCTTCATCTCTG 1180  
Db 1749 AGTGTCAATACCATTTTGCAGATACCTGGATGATGTTTCACTGAGTCACTTCCATCAG 1690  
Qy 1181 ATGTGGTGAACAATTCCTCTCGGCACTGGGAGGACCTTCCCGCCAGCCCTTGGTGGC 1240  
Db 1689 ATGCTGCAATGTACAACTCTGAAGCCCTGCCACCTCTCTTA----- 1644  
Qy 1241 CCCTTGGCCCACTCCCAACACTTCAGCAGCTTGGAGCTGGAGCCCGAGAGCCGAGCAGC 1300  
Db 1644 -----TGGCACCCCAACCTATGATC 1624  
Qy 1301 CCGTGGCAAGCCGAGGGAGCCGAGCCGCTCATCGGCTGGCTGGTGGCCATCA 1360  
Db 1623 CAATGCTTAAAGTTGATGACACACACTCGGATCCTGATTTGGCTGTGTGGCCATCA 1564  
Qy 1361 TCCTGCTCTGCTCATCATTTGCCCTCATGCTCTGGGGTGGAGTGGCGGAGGTTC 1420  
Db 1563 TCTTTATCTCTCGCCATCATTTGTATCATCTCTTGGAGGCGAGTTCTGGCAAAATGC 1504  
Qy 1421 TCAGCAAGCTGAAGGGGTGTTGGAAGAGGAGCTCAGCGTTTCACTCTCTGTCCTG 1480  
Db 1503 TGGAGAAGGCTTCTCGGAGGATGCTGGATGATGAATGACAGTCAAGCTTCCCTGCCAA 1444  
Qy 1481 GGGAGACTATCTCATCAACACCCCGCCAGGTCTTAGAGACCCCGCTAGCAGGAGC 1540  
Db 1443 GTGATCTAGCATGTTCAACAT-----AACCGCTCT 1411

Qy 1541 CCCGGCTCTGTGGGAATCCGGCCCACTCCGCTCCCTGTGTGCCCAATGGCTCTCGGTTC 1600  
Db 1410 CATCACTAGTGAACAAGGGTCCAACTGACTGACTTACGATCGCATCTTT----- 1363  
Qy 1601 TGCTCTCCAAATCCAGCTTACCGCTCTTCTTGGCCACTTACGCCGTCCCGCTCGAGGCC 1660  
Db 1363 -----CCCTTCCG 1354  
Qy 1661 CGGGCCCCCACCACCCGCTGGGCAAAACCCACCAACCCAGGCTTACAGTGGGACT 1720  
Db 1353 CTGACTACAGGAGCCATCCAGGCTGATACGAAACTCCAGAAATTTGCTCCAGGGAGG 1294  
Qy 1721 ATATGGAGCTGAGAAGCCAGCGCCCGCTTCTGCCCGCCAGCTTCCCGAGAACGCTTC 1780  
Db 1293 AGGAGTCAAGCTGCAAGCGTGTGTGAAGCCAGTCCAGCCAGTGGCCCTGAGGGGTGC 1234  
Qy 1781 CCATTTATGCCAGCTGACATTTTACCTCGAGGGGTTCACGGGGGTCAACACCTATG 1840  
Db 1233 CCCACTATGAGAGGCTGACATAGTGAACCTCAAGAGTACAGGAGGCAACATACT 1174  
Qy 1841 CTGTGCTGCACTGGCCCGCCAGGGGCGAGT---GGGATGGGCCCCCAGAGTGGATTTC 1897  
Db 1173 CAGTGGCTGGCTGCTGCTCAGGAAAGATGTGGCTGTGGAGGATTTCC 1114  
Qy 1898 CTCGATCTCGACTCCGCTTCAAGGAGAGCTTGGCGGGCCAGTTTGGGAGGTGAC 1957  
Db 1113 CCAGGAACTCTTAATTTTCAAGAGAGCTGGGAGAGGACAGTTTGGGAGGTTCATC 1054  
Qy 1958 TGTGTAGTCTCACAGCCCTCAAGATCTGGTCACTTTGATTTCCCTTAAATGTCGTA 2017  
Db 1053 TCTGTGAAGTGAAGGATGGAATAATCAAGACAAGATTTTCCCTAGATGTCACTG 994  
Qy 2018 AGGGACACCTTTGCTGTAGTGTCAAGATCTTACGCCAGATGCCACCAAGATGCCA 2077  
Db 993 CCAACCGAGCTGCTGTGGTGTGTAATAATGCTCCGAGCAGATGCCAACAAGATGCCA 934  
Qy 2078 GGAATGATTTCTGAAGAGTGAAGATCATGTCAGGCTCAAGGACCCCAACATCAITC 2137  
Db 933 GGAATGATTTCTTAAAGAGATAAGATCATGCTCGGCTCAAGGACCCCAACATCAITC 874  
Qy 2138 GGTGTGGTGGCTGTGTGTCAGGACGAGCCCTCTGTGATGATTTACTGACTACATGGAGA 2197  
Db 873 ATCTATTATCTGTGTATCACTGATGACCTCTCTGTATGATCACTGATACATGGAGA 814  
Qy 2198 AGGGGACCTCAACAGTTCCTCAGTGGCCACCGAGCTGGAGGACAGGAGCCGAGGGG 2257  
Db 813 ATGGAGATCTCAATCAGTTTCTTCCCGCCAGGAGCCCTAATTTCTTCTCCAGCG--- 756  
Qy 2258 CCCTTGGGAGCGGAGGCTGCGAGGGGCCACCATCAGCTACCCAAATGCTGCTGCATG 2317  
Db 756 -----ATGTACGCACTGTCAAGTTACACCAATCTGAAGTTTA 721  
Qy 2318 TGGCAGCCAGATCGCTCCGGATCGGCTATCTGGCCACACTCAACTTTGTACATCGGG 2377  
Db 720 TGGCTACCAATTTGCTCTGGCATGAAGTACTTCTCTCTTAATTTGTACCAG 661  
Qy 2378 ACTGCGCCAGCGGAATCGCTAGTTGGGAAATTTTACCATCAAAATCGCAGACTTTG 2437  
Db 660 ATCTGCGCACACGAAACTGTTAGTGGTGAAGAACTACACAATCAAGATAGCTGTTG 601  
Qy 2438 GCATGAGCGGAACCTCTATGCTGGGAGTATACCGTGTGCGAGGGCCGCGAGTGTGTC 2497  
Db 600 GAATGAGCAGGAACCTGTACAGTGGTGAATTTACCGGATCCAGGGCCGCGAGTGTCTCC 541  
Qy 2498 CCATCGCTGGATGGCTGGAGTGCATCTCATGGGAAGTTTACGACTGGGAGTGAGC 2557  
Db 540 CTATCGCTGGATGCTTGGAGAGTATCTTGTGGCAAGTTTCACTACAGCAAGTGATG 481  
Qy 2558 TGTGGGCTTTGTGTGACCCCTGTGGAGGTGCTGATGCTGTAGGCGCCAGCCCTTTG 2617  
Db 480 TGTGGGCTTTGGGGTACTTTTGGGAGACTTTCACCTTTTGTCAAGAACAGCCCTATT 421





```

; FILING DATE: 20-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/157563
; FILING DATE: 23-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: 854C1P1C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2820 bases
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-441-104A-4

```

Query Match 5.4%; Score 195.2; DB 4; Length 2820;  
 Best Local Similarity 59.6%; Pred. No. 1.5e-34;  
 Matches 392; Conservative 0; Mismatches 248; Indels 18; Gaps 3;

```

QY 1988 TCAGTCTTTGATTTCCCTTAAATGTGCGTAAGGACACACCTTTTCTGCTAGCTGTCAAGA 2047
DB 1920 TCCITGCTGAGTGCACACACCTCTGCTGAGCAGGACAGATCTGTGCTGTCAAG- 1979
QY 2048 TCTTACGCCAGATGCCACCAAGAAATGCCAGGAAATTTCTTCTGAAGAGGTGAAGATCA 2107
DB 1979 --GCACCTGAAGAGGCGTCCGAGAGTGTCTGCGGAGGCTTCAACGCTGAGCTGC 2036
QY 2108 TCTCAGCTCAAGGACCAACCAATCTTCGCTGCTGCGGCTGTGTGTCAGGACGACC 2167
DB 2037 TCACCATCTGCAGCAGCAGCAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2096
QY 2168 CCTCTGCTGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2227
DB 2097 CCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2156
QY 2228 ACCAGCTGGAGGACCAAGGACGCGGAGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2287
DB 2157 ATGAGCCTGATGCCAA-----GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2204
QY 2288 CCACCATCAGCTACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2347
DB 2205 GCGGCTGGGCTGGGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2264
QY 2348 ATCTGGCCACACTCAACTTTGTACATCGGACCTGGCCACGCGGAATGCTGCTGCTGCTGCTG 2407
DB 2265 ACCTGGCGGCTGCTATTTGTGACCGGCGACCTGGCCACACGCAACTGCTGCTGCTGCTGCTG 2324
QY 2408 AAAATTTTACCATCAAAATCGCAGACTTTGGATGAGCGCGGAACCTCTATGCTGGGACT 2467
DB 2325 AGGAGCTGGTGTGAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2384
QY 2468 ATTACCGTGTGCGGCGGCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2527
DB 2385 ATTACCGTGTGCGGCGGCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2444
QY 2528 TCATGGGAAGTTACGACTCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2587
DB 2445 TGTACCGTAAAGTTACACCGAGAGCGAGCTGTGGAGCTTCCGCGTGTGCTGCTGCTGCTGCTG 2504
QY 2588 TCTGATGCTCTGTAGGCGGCGGAGCTTTGGGAGCTTACCGAGCTACCGAGCAGGAGTCA 2645
DB 2505 TCTTACCTACGGCAAG---CAGCCCTGCTACCACTGCTTCAACACGAGGCAATCGA 2559

```

RESULT 10  
 US-08-440-816A-4  
 ; Sequence 4, Application US/08440816A

```

; Patent No. 5914237
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J.
; APPLICANT: Mark, Melanie R.
; APPLICANT: Sadick, Michael D.
; APPLICANT: Shelton, David L.
; APPLICANT: Wong, Wai Lee Tan
; TITLE OF INVENTION: KINASE RECEPTOR ACTIVATION ASSAY
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/440,816A
; FILING DATE: 15-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/286305
; FILING DATE: 05-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/170558
; FILING DATE: 20-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/157563
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: 854C1P1C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2820 bases
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-440-816A-4

```

Query Match 5.4%; Score 195.2; DB 4; Length 2820;  
 Best Local Similarity 59.6%; Pred. No. 1.5e-34;  
 Matches 392; Conservative 0; Mismatches 248; Indels 18; Gaps 3;

```

QY 1988 TCAGTCTTTGATTTCCCTTAAATGTGCGTAAGGACACACCTTTTCTGCTAGCTGTCAAGA 2047
DB 1920 TCCITGCTGAGTGCACACACCTCTGCTGAGCAGGACAGATCTGTGCTGTCAAG- 1979
QY 2048 TCTTACGCCAGATGCCACCAAGAAATGCCAGGAAATTTCTTCTGAAGAGGTGAAGATCA 2107
DB 1979 --GCACCTGAAGAGGCGTCCGAGAGTGTCTGCGGAGGCTTCAACGCTGAGCTGC 2036
QY 2108 TGTGCGGCTCAAGGACCCCAACATCTTCGCTGCTGCGGCTGTGTGTCAGGACGACC 2167
DB 2037 TCACCATCTGCAGCAGCAGCAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2096
QY 2168 CCTCTGCTGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2227
DB 2097 CCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2156
QY 2228 ACCAGCTGGAGGACCAAGGACGCGGAGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2287

```

Db 2157 ATGGACCTGATGCCAA-----GCTGCTGGCTGGTGGGGAGGATGTGGCTCCAG 2204  
Qy 2288 CCACCATCAGTACCCCAATGCTGCTGATGTCAGCGCCAGATCGCTCCGGCATCGCT 2347  
Db 2205 GCCCCTGGGTGGGGCAGCTGCTGGCCGTGCTAGCCAGGTGCTGGGGGATGGTGT 2264  
Qy 2348 ATCTGGCCACATCAACTTTGTACATCGGACCTGGCCAGCGGAACTGCTAGTTGGGG 2407  
Db 2265 ACCTGGGGGTCTGCAATTTGTGACCGGGACCTGGCCACACCAACTGCTAGTGGGCC 2324  
Qy 2408 AAAATTTCCACATCAAAATCCGAGACTTTGGCATGAGCGGAACTTATGCTGGGGACT 2467  
Db 2325 AGGACTGGTGGTCAAGATTGGTGAATTTGGCATGAGCAGGATATCTACAGCCCGAT 2384  
Qy 2468 ATACCGTGTGAGGGCGGGCAGTGTGCTCCATCGCTGGATGGCTGGAGTGCATCC 2527  
Db 2385 ATTACCGTGTGGAGGGCGCCACCATGCTGCCATTCGCTGGATGCCCGGAGATCC 2444  
Qy 2528 TCATGGGAAGTTTCAGCACTTCCAGTGTGAGTGTGGGCTTTGTGTGACCCCTGTGGGAGG 2587  
Db 2445 TGTACCTAAGTTTCAACACCGAGAGCGACGTGTGGAGCTTGGCGGTGGTCTGTGGAGA 2504  
Qy 2588 TGTGTGATGCTGTGAGGGCGGACCTTTTGGCGAGCTCACCGACGAGAGGTATCGA 2645  
Db 2505 TCTTCACTAGCGCAAG---CAGCCCTGTGTACAGCTCTCCAAACAGGAGGCAATCGA 2559

RESULT 11  
US-08-306-691B-23  
; Sequence 23, Application US/08306691B  
; Patent No. 5734039  
; GENERAL INFORMATION:  
; APPLICANT: Calabretta, Bruno  
; APPLICANT: Skorski, Tomasz  
; TITLE OF INVENTION: ANTISENSE  
; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES  
; NUMBER OF SEQUENCES: 55  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seidel, Gonda, Lavorina & Monaco, P.C.  
; STREET: Two Penn Center, Suite 1800  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: U.S.A.  
; ZIP: 19102

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 720 kb  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: WordPerfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/306,691B  
; FILING DATE: September 15, 1994  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Monaco, Daniel A.  
; REGISTRATION NUMBER: 30,480  
; REFERENCE/DOCKET NUMBER: 8321-8  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 568-8383  
; TELEFAX: (215) 568-5549  
; TELEX: NO. 5734039e  
; INFORMATION FOR SEQ ID NO: 23:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2301 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
US-08-306-691B-23

Query Match 5.3%; Score 193.6; DB 2; Length 2301;  
Best Local Similarity 59.4%; Pred. No. 3.1e-34;  
Matches 391; Conservative 0; Mismatches 249; Indels 18; Gaps 3;  
Qy 1988 TCAGTCTTGATTTCCCTTAAATGTGGTAAGGACACCTTTCTGCTGGTGTCTCAAGA 2047  
Db 1275 TCCTTGTGTAGTGGCACAACTCTGCTGAGCAGGACAGATGCTGTGCTGTCTCAAG- 1334  
Qy 2048 TCTTACGGCCAGATGCCACCAAGAAATCCAGGAATGCTCGGAGAGTCTTCCAACTGAGGCTGAGTGC 2107  
Db 1334 --GCACTGAAGAGGCGTCCGAGAGTCTCGGAGGACTTCCAACTGAGGCTGAGTGC 1391  
Qy 2108 TGTGAGGCTCAAGAGACCCAAACATCATTCGGCTGCTGGCGTGTGTGTCAGGACGAC 2167  
Db 1392 TCACCATCTGCAGCAGCAGCAGCATGCTGCGCTTCTTCGGCGTCTGCACGAGGCGCC 1451  
Qy 2168 CCTCTGCACTTACTGACTACATGAGAGAACGGGACCTCAACCACTTCTCTAGTGGCC 2227  
Db 1452 CCCTGCTCATGCTTCTGAGTATATGCGGACGCGGAGCTCAACCGCTTCTCTCGATCC 1511  
Qy 2228 ACCAGTGGAGGACAAAGCAGCAGCGGCGGCGCTGGGAGCGGCGAGCTGCGGAGGGC 2287  
Db 1512 ATGACCCGATGCCAA-----GCTGCTGGTGTGGGAGGATGTGCTCCAG 1559  
Qy 2288 CCACCATCAGTACCCCAATGCTGCTGTCATGTGGCAGCGCCAGATCGCTCCGGCATGCGCT 2347  
Db 1560 GCGCCTGGGTCTGGGCGAGCTGCTGGCGGTGGCTAGCCAGGTGCTGCGGGATGCTGT 1619  
Qy 2348 ATCTGGCCACACTCAACTTTGATCATCGGACCTGGCCAGCGGAACTGCTAGTGGGG 2407  
Db 1620 ACCTGGCGGTCTGCTCATTTTGTGACCGGAGCTGGCGACACCACTGCTCTAGTGGGC 1679  
Qy 2408 AAAATTTCAACCATCAAAATCCAGACTTTGTCATGAGCGGAACTCTATCTGCTGGGACT 2467  
Db 1680 AGGACTGCTGCTCAAGATGCTGTTGGTATGTCATGAGCGGAGATCTACAGCAGGACT 1739  
Qy 2468 ATTACCGTGTGAGGGCGGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2527  
Db 1740 ATTACCGTGTGAGGGCGGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1799  
Qy 2528 TCATGGGAAGTTTCAGCACTTCCAGTGTGAGTGTGGGCTTTGTGTGACCCCTGTGGGAGG 2587  
Db 1800 TGTACCGTAAATTCCACCAGGAGAGCGAGTGTGGAGCTTGGCGGTGCTGCTGCGAGA 1859  
Qy 2588 TGTGTGATGCTGTGAGGGCGGCGAGCTTGGGAGCTTCCAGGAGTGTGCTGCTGCTGCTGCT 2645  
Db 1860 TCTTCACTAGCGCAAG---CAGCCCTGTTACAGCTCTCCAAACAGGAGGCAATCGA 1914

RESULT 12  
PCT-US93-06251-78  
; Sequence 78, Application PC/TUS9306251  
; GENERAL INFORMATION:  
; APPLICANT: Wickstrom, Eric and Rife, Jason P.  
; TITLE OF INVENTION: Trivalent Synthesis of Oligonucleotides Containing  
; TITLE OF INVENTION: Stereospecific Alkylphosphonates and Arylphosphonates  
; NUMBER OF SEQUENCES: 93  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/06251  
; FILING DATE: 19930630  
; CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:  
NAME: Digilio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 8586  
TELEPHONE: 516-742-4343  
TELEFAX: 516-742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 78:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2301 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
PCT-US93-06251-78

Query Match 5.3%; Score 193.6; DB 5; Length 2301;  
Best Local Similarity 59.4%; Pred. No. 3.1e-34;  
Matches 391; Conservative 0; Mismatches 249; Indels 18; Gaps 3;  
QY 1988 TCAGTCTTGAATTCCTCCCTTAATGTGCTGAAGGACACCCCTTTCTGCTAGCTGTCAAGA 2047  
Db 1275 TCCCTTGGCTGAGTGCCACCAACCTCTGCTGAGCAGGACAGATGCTGGTGGCTGTCAAG- 1334  
QY 2048 TCTTACGGCCAGATGCCACCAAGAAATGCCAGGAATGATTTCTGAAAGAGGTGAAGATCA 2107  
Db 1334 --GCACTAAGGAGCGTCCGAGAGTGTCTCGGAGGACTTCCACGTGAGCTGAGCTGC 1391  
QY 2108 TGTGAGGCTCAAGGACCCCAACATCATTCGGCTGCTGGGGTGTGTGTGAGGAGCAC 2167  
Db 1392 TCACCATGCTGCAGCACCAGCACATCTGCTGCTTCTTGGGGTGTGCACCGAGGGCGCC 1451  
QY 2168 CCCTCTGATGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2227  
Db 1452 CCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1511  
QY 2228 ACCAGCTGGAGGACAGGACGCGAGGGGGCCCTGGGGAGCGGACGCTGCGCAGGGGC 2287  
Db 1512 ATGGACCGATGCCAA-----GCTGCTGGCTGGTGGGAGGATGGCTCCAG 1559  
QY 2288 CCACCATCAGCTACCAATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2347  
Db 1560 GCCCCTGGTCTGGGGCAGCTGCTGGCCGTGGCTAGCCAGGCTGCTGCGGGATGGTGT 1619  
QY 2348 ATCTGGCCACACTCAACTTTGTACATCGGACCTGGCCACGCGAATGCTAGTTGGG 2407  
Db 1620 ACCTGGCGGGTCTGCAATTTGTGACCGGGACCTGGCCACGCAACTGTCTAGTGGGC 1679  
QY 2408 AAAATTTCCACCATCAAAATCGCAGACTTTGGCATGAGCGCGAACTCTATGCTGGGACT 2467  
Db 1680 AGGAGTGGTGTCAAGATTTGGTGAATTTGGCATGAGCAGGAGATATCTACAGCAGCACT 1739  
QY 2468 ATTACCGTGTGAGGGCGGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2527  
Db 1740 ATTACCGTGTGAGGGCGGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1799  
QY 2528 TCATGGGAAAGTTACAGACTCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2587  
Db 1800 TGATCGGTAAAGTTACACCGAGAGCGAGTGTGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1859  
QY 2588 TGCTGATGCTGTGAGGGCCAGCCCTTTGGGAGCTCACCGAGCAGCAGTGTATCGA 2645  
Db 1860 TCTTCACTACGGCAAG---CAGCCCTGGTACCAGCTCTCCAAACAGGAGCAATCGA 1914

RESULT 13  
US-08-271-454-1  
; Sequence 1, Application US/08271454  
; Patent No. 5601820  
; GENERAL INFORMATION:  
; APPLICANT: Brodeur, Garrett M.

APPLICANT: Nakagawara, Akira  
TITLE OF INVENTION: Compositions and Methods of Making and  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5601820rls  
STREET: One Liberty Place, 46th floor  
CITY: Philadelphia  
STATE: PA  
COUNTRY: USA  
ZIP: 19103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/271,454  
FILING DATE: 07-JUL-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: DeLuca, Mark  
REGISTRATION NUMBER: 33,229  
REFERENCE/DOCKET NUMBER: CH-0535  
TELEPHONE: (215) 568-3100  
TELEFAX: (215) 568-3439  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3707 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: both  
MOLECULE TYPE: CDNA  
US-08-271-454-1

Query Match 5.2%; Score 190; DB 1; Length 3707;  
Best Local Similarity 60.7%; Pred. No. 2.3e-33;  
Matches 357; Conservative 0; Mismatches 210; Indels 21; Gaps 2;

QY 2059 GATGCCACCAAGATGCCAGGAATGATTTCTGAAAGAGTGAAGATCATGTCGAGGCTC 2118  
Db 2077 GATGCCAGTGAATGACGACGCAAGGACTTCCACCGTGGAGCGGAGCTCTGACCAACCTC 2136  
QY 2119 AAGGCCCAACATCATTCGGCTGCTGGCGGTGTGTGTCAGGAGCAGCCCTCTGCGATG 2178  
Db 2137 CAGCATGAGCAGATCGTCAAGTTCTATGGCGTCTGCGTGGAGGGCGCCCTCTCATATG 2196  
QY 2179 ATTACTGACTAGATGAGAAACGCGACCTCAACAGTTCTCTCAGTCCACACAGCTGGAG 2238  
Db 2197 GTCTTTGACTAGATGAGCATGGGACCTCAACAAGTTCTCAGGGCAGACGGCCCTGAT 2256  
QY 2239 GACAAGCAGCAGGGGGCCCTTGGGACGCGGAGGTGCGCAGGGGCGCCACCATCAGC 2298  
Db 2257 GCGTGTCTGATGGTGGGCAACCGCCGCGGAACTGACGCGAGTGC----- 2306  
QY 2299 TACCCAACTGCTGCTGATGTCGACGCGGAACTGCTAGTTGGGAAATTTTACC 2358  
Db 2306 -----AGATGCTGCATATAGCCAGAGATCGCGCGGGCATGGTCTACCTGGCTCC 2358  
QY 2359 CTCAACTTTGATCGGACCTGCGGACGCGGAACTGCTAGTTGGGAAATTTTACC 2418  
Db 2359 CAGCACTTCTGTCACCGGATTTGGCCACGAGAACTGCTGCTGGGAGAACTTGTG 2418  
QY 2419 ATCAAATCGAGACTTTGGCATGAGCCGGAACCTCTTATGTGGGAGTATTACCGTGTG 2478  
Db 2419 GTGAAATCGGGAGCTTTGGGATGTCCCGGAGCTGTACAGCACTGACTTACAGGGTC 2478  
QY 2479 CAGGGCGGGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2538  
Db 2479 GGTGGCCACACAACTGCTGCCATTCGCTGGATGCTCCAGAGAGCATGATGTACAGGAA 2538

2539	TTACACACTGCAGTAGTACGCTGTGGCCCTTTGGTGTGACCCCTGTGGGAGTGCTGTGATGCTC	2598
QY		
2539	TTACACACGGAAGGACGCTCTGGACCTGGGGGTCGTGTTGTGGGAGATTTCACCC	2599
Db		
2599	TGTAGGCCCCAGCCCTTTGGGCAGCTCACCGACGACGAGGTCAATCGAG	2646
QY		
2596	TATGGCAACACGCCCTGTTACCACTGTCAACAATGAGGTGATGAG	2643
Db		

```

RESULT 14
US-08-286-305A-6
; Sequence 6, Application US/08286305A
; Patent No. 5766863
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J.
; APPLICANT: Mark, Melanie R.
; APPLICANT: Sadick, Michael D.
; APPLICANT: Shelton, David L.
; APPLICANT: Wong, Wai Lee Tan
; TITLE OF INVENTION: KINASE RECEPTOR ACTIVATION ASSAY
; NUMBER OF SEQUENCES: 11

```

Db	2227	GTCTTTTGATGATGAAGCATGGGACCTCAACAAAGTTCTTCAGGGCACACGCGCCCTGAT	2286
QY	2239	GACAAGGCAGCGAGGGGGCCCTGGGACAGGGCAGGCTGCGCAGGGGGCCACCATCAGC	2298
Db	2287	GCCGTGCTATGGCTGAGGGNACCCGCCACGGNACTGACGCACTGCG	2336
QY	2299	TACCCAAATGCTGCTGCATGTGGCAGCCAGATCGCCTCGGSCATGCGCTATCTGGCCACA	2358
Db	2336	-----AGATGCTGCATATAGCCACGACAGATCGCGCGGSCATGTCTACCTGGCGTCC	2388
QY	2359	CTCAACTTTGTACATCGGGACCTGGCCACGCGGAACCTAGTCTGGGGGAAAATTCACC	2418
Db	2389	CAGCACTTGTGCACCGGGATTTGCCACACGAGAACTCCCTGGTCGGGGGAGAACTTGCTG	2448
QY	2419	ATCAAAATCGCAGACTTTGGCATGAGCGGGAACCTCTATGCTGGGGACTATTACCGTGTC	2478
Db	2449	GTGAAATTCGGGGACTTTGGGATGTCCCGGACGTGTACAGCACTGACTACTACAGGTC	2508
QY	2479	CAGGGCCGGGCAGTGTGCCATCCGCTGGATGGCTGGGAGTGCATCTCATGGGGGAG	2538
Db	2509	GGTGGCCACACAATGCTGCCCATTCGCTTGGATGCTTCAGAGAGCATCTGTACAGGAA	2568
QY	2539	TTACAGACTGCGAGTGACGTGTGGGCCCTTGGTGTGACCCCTGTGGGAGGTGCTGATGCTC	2598
Db	2569	TTACAGACGAAAGCGACGCTGGAGCCTGGGGCTGCTGTTGTGGGAGATTTCAC	2625
QY	2599	TGTAGGGCCCCGCCCTTTGGGCAAGCTCACCGACGAGCGTCACTGAG	2646
Db	2626	TATGGCAACAGCCCTGTTACCACTGTCAGCACTGTCAACCAATGAGGTGATAGAG	2673

RESULT 15

US-08-359-705B-1  
 ; Sequence 1, Application US/08359705B  
 ; Patent No. 5844092  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Presta, Leonard G.  
 ; APPLICANT: Shelton, David L.  
 ; APPLICANT: Urfer, Roman  
 ; TITLE OF INVENTION: Human trk Receptors and Neurotrophic Factor Inhibitors  
 ; NUMBER OF SEQUENCES: 41  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Genentech, Inc.  
 ; STREET: 1 DNA Way  
 ; CITY: South San Francisco  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 94080  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: WinPatIn (Genentech)  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/359,705B  
 ; FILING DATE: 20-Dec-1994  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/286846  
 ; FILING DATE: 08/10/94  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/215139  
 ; FILING DATE: 03/18/94  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Torchia, Phd., Timothy E.  
 ; REGISTRATION NUMBER: 36,700  
 ; REFERENCE/DOCKET NUMBER: P0873P2  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 650/225-8674  
 ; TELEFAX: 650/952-9881  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:

Search completed: November 4, 1999, 07:29:20  
Job time: 6308 sec

Human trkC receptor  
Potential tyrosine

## ALIGNMENTS

44	123.8	3.4	2712	1	T00690
45	121	3.3	755	1	Q30725

## RESULT 1

T93785 standard; cDNA: 3962 BP.

16-FEB-1998 (first entry)

Human mammary calcitonin receptor kinase 10 (MCK-10) cDNA sequence.

proliferative disease; cancer; insulin receptor family;

## 15339 /

(ADVEV) ALVES F H E.  
(ULLR) ULLRICH A.  
Alves FHE, Ullrich A;  
WPI: 97-511869/47.  
P-PSDB: W34672.  
Truncated receptor tyrosine kinase CKK-2 - and nucleic acid coding for it, useful for cancer diagnosis  
Disclosure; Fig 1; 70pp; English.  
The present sequence represents the cDNA of a mammary carcinoma kinase, called MCK-10. This kinase belongs to a novel family of receptor tyrosine kinases, and expression is associated with proliferative diseases such as cancer. The MCK-10 receptor tyrosine kinase has extensive sequence similarity to the insulin receptor family. The MCK-10 gene was isolated by PCR using 2 degenerate oligonucleotide primer pools, using a template cDNA synthesised by reverse transcription of poly-A RNA from the human mammary carcinoma cell line MCF7. MCK-10 is expressed in brain tissue, and the protein shares homology with the tyrosine kinase neurotrophin receptor. Modulation of MCK-10 activity therefore may be used for treatment of neurological disorders. MCK-10 is also expressed in a variety of cancer cell lines and tumour tissue. The present sequence, or parts of it, can be used for diagnostic purposes to detect aberrant expression of MCK-10 genes. Inhibitors of MCK-10 receptor activity may have therapeutic value in the treatment of diseases such as cancer.  
Sequence 3962 BP: 735 A: 1234 C: 1182 G: 811 T:

Query Match 94.9%; Score 3451; DB 1; Length 3962;  
Best Local Similarity 97.0%; Pred. No. 0;  
Matches 3589; Conservative 0; Mismatches 5; Indels 105; Gaps

17 GTTGGACTTGAAGGAATGCCAAGAGATGCTGCCCCACCCCTTAGGCCCGGAGGATCAG 76  
|||||  
256 GTTGGACTTGAAGGAATGCCAAGAGATGCTGCCCCACCCCTTAGGCCCGGAGGATCAG 315  
|||||  
77 GAGCTATGGGACCAGAGGGCCCTGTCACTTTACTGCTGCTGCTCTTTGTTGGCAAGTGGAG 136  
|||||  
316 GAGCTATGGGACCAGAGGGCCCTGTCACTTTACTGCTGCTGCTCTTTGTTGGCAAGTGGAG 375  
|||||  
137 ATGCTGACATGAAGGACACATTTTGATCTGCCAAGTGGCCGTATGCCCTGGGCATGTCAGG 196  
|||||  
376 ATGCTGACATGAAGGACATTTTGATCTGCCAAGTGGCCGTATGCCCTGGGCATGTCAGG 435  
|||||  
197 ACCGGACCATCCAGACATGACATCTCTGCTTCCAGTTCCTTGGTCAGATTCCACTGCGCG 256  
|||||  
436 ACCGGACCATCCAGACATGACATCTCTGCTTCCAGTTCCTTGGTCAGATTCCACTGCGCG 495  
|||||  
257 CCCGCCACACAGCTTGGAGAGCAGTGCAGGGGATGGGGCCCTGGTGCCTCCCGCAGGGTCCG 316

Db 496 CCCGCCACAGAGTTGGAGAGCAGTACCGGGGATGGGGCTTGGTGGCCCGCAGGCTCGG 555  
Qy 317 TGTTCCTCCAAAGAGAGAGTACTTGCAGGTGGATCTACAAACAGCTGCACCTGGTGGCTC 376  
Db 556 TGTTCCTCCAAAGAGAGAGTACTTGCAGGTGGATCTACAAACAGCTGCACCTGGTGGCTC 615  
Qy 377 TGGTGGGACCCAGGACGGCATCCGGGGGCGCTGGGCAAGGAGTTCTCCCGGAGCTACC 436  
Db 616 TGGTGGGACCCAGGACGGCATCCGGGGGCGCTGGGCAAGGAGTTCTCCCGGAGCTACC 675  
Qy 437 GGCTGGCTTACTCCCGGATGGTCCGCTGGATGGCTGGAAGGACCGCTGGGGTCAGG 496  
Db 676 GGCTGGCTTACTCCCGGATGGTCCGCTGGATGGCTGGAAGGACCGCTGGGGTCAGG 735  
Qy 497 AGGTGATCTCAGGCAATGAGGACCTCAGGAGTGGTCTGAAGGACCTTGGGCCCCCA 556  
Db 736 AGGTGATCTCAGGCAATGAGGACCTCAGGAGTGGTCTGAAGGACCTTGGGCCCCCA 795  
Qy 557 TGGTTGCCGACTGGTTCGCTTACCCCGGGCTGACCGGGTCATGAGGCTCTGTCTGC 616  
Db 796 TGGTTGCCGACTGGTTCGCTTACCCCGGGCTGACCGGGTCATGAGTGTCTGTCTGC 855  
Qy 617 GGATAGACTCTATGGCTGCTCTGGAGGATGGACTCTCTTACACCGCCCTGTGG 876  
Db 856 GGGTAGACTCTATGGCTGCTCTGGAGGATGGACTCTCTTACACCGCCCTGTGG 915  
Qy 677 GGCAGACAATGATTATCTGAGCGCTGTACCTCAACGACTCCACCTATGACGGACATA 736  
Db 916 GGCAGACAATGATTATCTGAGCGCTGTACCTCAACGACTCCACCTATGACGGACATA 975  
Qy 737 CCGTGGGCGGACTGCAGTATGGGGTCTGGGCCAGCTGGCAGATGGTGGTGGGGCTGG 796  
Db 976 CCGTGGGCGGACTGCAGTATGGGGTCTGGGCCAGCTGGCAGATGGTGGTGGGGCTGG 1035  
Qy 797 ATGACTTTAGGAAGAGTCAGGAGCTGGGGTCTGGCCAGGCTATGAGTATGGGATGGA 856  
Db 1036 ATGACTTTAGGAAGAGTCAGGAGCTGGGGTCTGGCCAGGCTATGAGTATGGGATGGA 1095  
Qy 857 GCAACACAGCTTCTCAGTGGCTATGTGGAGATGGAGTTGAGTTGACCGCTGAGGG 916  
Db 1096 GCAACACAGCTTCTCAGTGGCTATGTGGAGATGGAGTTGAGTTGACCGCTGAGGG 1155  
Qy 917 CTTTCCAGGCTATGCAGTCCACTGTAAACAAATGCACAGCTGGGAGCCCGTCTGCGTG 976  
Db 1156 CTTTCCAGGCTATGCAGTCCACTGTAAACAAATGCACAGCTGGGAGCCCGTCTGCGTG 1215  
Qy 977 GCGGGGTGGAATGCTCTCCGGCTGGCCCTGCCATGGCTGGGAGGGGAGCCCATGC 1036  
Db 1216 GCGGGGTGGAATGCTCTCCGGCTGGCCCTGCCATGGCTGGGAGGGGAGCCCATGC 1275  
Qy 1037 GCAACACTAGGGGGCAACCTGGGGACCCAGAGCCCGGGCTGTCTCAGTGGCCCTTG 1096  
Db 1276 GCAACACTAGGGGGCAACCTGGGGACCCAGAGCCCGGGCTGTCTCAGTGGCCCTTG 1335  
Qy 1097 GCGGCGGTGTGGCTGCTTCTGCAAGTCCGCTTCTCTTTGGGGGCGCTGGTTACTCT 1156  
Db 1336 GCGGCGGTGTGGCTGCTTCTGCAAGTCCGCTTCTCTTTGGGGGCGCTGGTTACTCT 1395  
Qy 1157 TCAGGGAATCTCTTCACTCTGATGTGGTGAACAATTCCTCTCCGGCACTGGGAGCA 1216  
Db 1396 TCAGGGAATCTCTTCACTCTGATGTGGTGAACAATTCCTCTCCGGCACTGGGAGCA 1455  
Qy 1217 CTTTCCCGCAGCCCTTGGTGGCGCTGGCCACCTCCCACTCCCACTTCAGAGCTTGG 1276  
Db 1456 CTTTCCCGCAGCCCTTGGTGGCGCTGGCCACCTCCCACTTCAGAGCTTGG 1515  
Qy 1277 AGCTGGAGCCAGAGCCAGAGCCCGTGGCCAAAGCCGAGGGAGCCGACGCCATCC 1336  
Db 1516 AGCTGGAGCCAGAGCCAGAGCCCGTGGCCAAAGCCGAGGGAGCCGACGCCATCC 1575  
Qy 1337 TCATGGCTGCTGGTGGCCATCTCTGCTCTGCTCTATCTATCTGCTCTGCTCT 1396  
Db 1576 TCATGGCTGCTGGTGGCCATCTCTGCTCTGCTCTATCTATCTGCTCTGCTCT 1635

Qy 1397 GCGGCTGCACTGGCGCAGGCTCTCTCAGCAAGGCTGAACGAGGCTGTGGAAGAGGAGC 1456  
Db 1636 GCGGCTGCACTGGCGCAGGCTCTCTCAGCAAGGCTGAACGAGGCTGTGGAAGAGGAGC 1695  
Qy 1457 TGACGGTTACCT 1516  
Db 1696 TGACGGTTACCT 1755  
Qy 1517 GAGAGCCACCCCGTACAGGAGCCCGGCTCTGTTGGAATCCCGCCACTCCGCTCCCT 1576  
Db 1756 GAGAGCCACCCCGTACAGGAGCCCGGCTCTGTTGGAATCCCGCCACTCCGCTCCCT 1815  
Qy 1577 GTGTCTCCCAATGGCT 1636  
Db 1816 GTGTCTCCCAATGGCT 1875  
Qy 1637 CTTAGCGCGCTCCCGCTCGAGGCGCGGCGCCCGCCACACCGCTGGGCCAAACCCACCA 1696  
Db 1876 CTTAGCGCGCTCCCGCTCGAGGCGCGGCGCCCGCCACACCGCTGGGCCAAACCCACCA 1935  
Qy 1697 ACACCCAGGCTTACAGTGGGACTATATGAGCTCTGAGAGCTGAGAGGAGGCGCCCGCTCTGC 1756  
Db 1936 ACACCCAGGCTTACAGTGGGACTATATGAGCTCTGAGAGCTGAGAGGAGGCGCCCGCTCTGC 1995  
Qy 1757 CCCCACCTCCCGCAGAACAGCTTCCCGCATATATGCGGAGCTGACATTTGTTACCTGCGAG 1816  
Db 1996 CCCCACCTCCCGCAGAACAGCTTCCCGCATATATGCGGAGCTGACATTTGTTACCTGCGAG 2055  
Qy 1817 GCGTCACCGGGGCAACACCTATGCTGTGCTCTGCACTCCCGCCAGGGGCACTCGGGGATG 1876  
Db 2056 GCGTCACCGGGGCAACACCTATGCTGTGCTCTGCACTCCCGCCAGGGGCACTCGGGGATG 2115  
Qy 1877 GCGCCCCAGAGTGGATTTCCCTCGATCTCGACTCTCAAGGAGAGCTTGGCGAGG 1936  
Db 2116 GCGCCCCAGAGTGGATTTCCCTCGATCTCGACTCTCAAGGAGAGCTTGGCGAGG 2175  
Qy 1937 GCGAGTTTGGGAGGTGCACCTGTGTGAGGTGCAGAGCCCTCAAGATCTGGTCACTGTG 1996  
Db 2176 GCGAGTTTGGGAGGTGCACCTGTGTGAGGTGCAGAGCCCTCAAGATCTGGTCACTGTG 2235  
Qy 1997 ATTTCCTCTTAAATGTGGTAAAGGACACCTTTGCTGGTAGCTGTCAAGATCTTACGGC 2056  
Db 2236 ATTTCCTCTTAAATGTGGTAAAGGACACCTTTGCTGGTAGCTGTCAAGATCTTACGGC 2295  
Qy 2057 CAGATGCCACCAAGATG-----CCAGGAATGATTTCTCTGAAAGAGG 2098  
Db 2296 CAGATGCCACCAAGATGCGAGCTTCTCTTGTCTCCAGGAATGATTTCTCTGAAAGAGG 2355  
Qy 2099 TGAAGATCATGTGAGGCTCAAGGACCCCAACATCATTCGGCTGCTGGCGGTGTGTGTC 2158  
Db 2356 TGAAGATCATGTGAGGCTCAAGGACCCCAACATCATTCGGCTGCTGGCGGTGTGTGTC 2415  
Qy 2159 AGGAGACCCCTCTGCAATGATTTACTGATGATGAGGAGAACCGGACCTCAACAGTTCC 2218  
Db 2416 AGGAGACCCCTCTGCAATGATTTACTGATGATGAGGAGAACCGGACCTCAACAGTTCC 2475  
Qy 2219 TCAGTGGCCACAGCTGGAGGACAAAGGAGCGGCGGCGCCCTGGGAGCGGAGGCTG 2278  
Db 2476 TCAGTGGCCACAGCTGGAGGACAAAGGAGCGGCGGCGCCCTGGGAGCGGAGGCTG 2535  
Qy 2279 GCGAGGGGCGCCACCATCAGCTACCCAATGCTGCTGATGTGGCAGCCAGATCGGCTCCG 2338  
Db 2536 GCGAGGGGCGCCACCATCAGCTACCCAATGCTGCTGATGTGGCAGCCAGATCGGCTCCG 2595  
Qy 2339 GCATGGCTTATCTGGCAGACTCAACTTTGTACATCGGGAGCTGGCCAGCGGGAATGCC 2398  
Db 2596 GCATGGCTTATCTGGCAGACTCAACTTTGTACATCGGGAGCTGGCCAGCGGGAATGCC 2655  
Qy 2399 TAGTTGGGAAAAATTTACCATCAAAATCGCAGACTTTGGCATGAGCGGGAACCTCTATG 2458  
Db 2656 TAGTTGGGAAAAATTTACCATCAAAATCGCAGACTTTGGCATGAGCGGGAACCTCTATG 2715



QY 2459 CTGGGAGCTATACCGTGTGAGGCGCGGCGAGTGTGCGCCATCCGCTGGATGCCCTGGG 2518  
DB 2716 CTGGGAGCTATACCGTGTGAGGCGCGGCGAGTGTGCGCCATCCGCTGGATGCCCTGGG 2775  
QY 2519 AGTGCATCTCATGGGAAGTTACAGACTGCAGTGCAGTGTGCGCTTTGGTGTGACCC 2578  
DB 2776 AGTGCATCTCATGGGAAGTTACAGACTGCAGTGCAGTGTGCGCTTTGGTGTGACCC 2835  
QY 2579 TGTGGAGGTGCTGATGCTCTGTAGGCGCCAGCCCTTTGGGAGCTCACCGACGAGCAGG 2638  
DB 2836 TGTGGAGGTGCTGATGCTCTGTAGGCGCCAGCCCTTTGGGAGCTCACCGACGAGCAGG 2895  
QY 2639 TCATCGAAGCGGGGGAGTCTTCGCGGACACAGGCGCGGCGAGTGTACCTGTGCCGCGC 2698  
DB 2896 TCATCGAAGCGGGGGAGTCTTCGCGGACACAGGCGCGGCGAGTGTACCTGTGCCGCGC 2955  
QY 2699 CGCCTGCTGCCGCGAGGCGCTATATGAGCTGATGCTTCGGTGTGAGCGGGAGTCTG 2758  
DB 2956 CGCCTGCTGCCGCGAGGCGCTATATGAGCTGATGCTTCGGTGTGAGCGGGAGTCTG 3015  
QY 2759 AGCAGCGACCACTCTTTCCAGCTGCTATGCTGCTTCGGGAGGATGCACCTCAACACGG 2818  
DB 3016 AGCAGCGACCACTCTTTCCAGCTGCTATGCTGCTTCGGGAGGATGCACCTCAACACGG 3075  
QY 2819 TGTGAATCACATCCAGCTGCCCTCCCTCAGGAGTGTATCCAGGGAAGCCAGTGCACA 2878  
DB 3076 TGTGAATCACATCCAGCTGCCCTCCCTCAGGAGTGTATCCAGGGAAGCCAGTGCACA 3135  
QY 2879 CTAAACAGAGGAGACAAATGCGACCTCTGCCCTTCCCTCCCGACAGCCCATCCTCT 2938  
DB 3136 CTAAACAGAGGAGACAAATGCGACCTCTGCCCTTCCCTCCCGACAGCCCATCCTCT 3195  
QY 2939 AATAGAGCAGTGTGAGCTGC----- 2959  
DB 3196 AATAGAGCAGTGTGAGCTGCAGTGGCTGGGCGCCACCCAGGAGCTGATGCCCTTCTC 3255  
QY 2959 -----AGAGCCCGCTGTGC 2972  
DB 3256 CCTCTCTGAGACACACTCATATGCTCCCTCTCTCTTCTCTCTCTCTCTCTCTCTCTG 3315  
QY 2973 CCCACCCAGCTGCTCTGTGATGGATCTCCCTCCAGCTCTCTCTCTCTCTCTCTCTCTG 3032  
DB 3316 CCCACCCAGCTGCTCTGTGATGGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 3075  
QY 3033 AAGGTGGGAGAAATATAGATAGACACTGGACATGGCCATGGCCATGGGAGCTGGGCCCC 3092  
DB 3376 AAGGTGGGAGAAATATAGATAGACACTGGACATGGCCATGGGAGCTGGGCCCC 3435  
QY 3093 ACTGACAACTGATCTCTGGAGAGTGTGCTGGCGCCCGCCAGCTTCTCTCTCTCTCTCTC 3152  
DB 3436 ACTGACAACTGATCTCTGGAGAGTGTGCTGGCGCCCGCCAGCTTCTCTCTCTCTCTCTC 3494  
QY 3153 ACATGGACCCACTGGCTGAGATCTGGGGTGGAGGAGACAGAGGAGAGGAGAAATG 3212  
DB 3495 ACATGGACCCACTGGCTGAGATCTGGGGTGGAGGAGACAGAGGAGAGGAGAAATG 3554  
QY 3213 TTTCTTGTGCTGCTCTCTGATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3272  
DB 3555 TTTCTTGTGCTGCTCTCTGATCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3614  
QY 3273 GAACACTGGACCTGGGGGTAGCGCGCCCGCCAGCTTCTCTCTCTCTCTCTCTCTCTCTCT 3332  
DB 3615 GAACACTGGACCTGGGGGTAGCGCGCCCGCCAGCTTCTCTCTCTCTCTCTCTCTCTCTCT 3674  
QY 3333 CAGCTTCTGATGACTCTCTTCTAGCCTATAGCTTCTGTGGAGTAAATATTTGGGATT 3392  
DB 3675 CAGCTTCTGATGACTCTCTTCTAGCCTATAGCTTCTGTGGAGTAAATATTTGGGATT 3734  
QY 3393 GGGGGAAGAGGAGGAGCAACGCCCATAGCTTGGGGTGGAGATCTCTCTCTCTCTCTCTCTCT 3452  
DB 3735 GGGGGAAGAGGAGGAGCAACGCCCATAGCTTGGGGTGGAGATCTCTCTCTCTCTCTCTCT 3794  
QY 3453 CACATTGATTTTCTATTAATCACTTGGGGTTGTACATTTTGGGGGGGAGAGACACAGAT 3512

DB 3795 CACATTGATTTTCTATATCATCTTGGGTTTGTACATTTTGGGGGAGAGACACAGAT 3854  
QY 3513 TTTTACACTAATATATGACCTAGCTTGGGCAATTTAAATCCCTGCTACTAGGAGGTA 3572  
DB 3855 TTTTACACTAATATATGACCTAGCTTGGGCAATTTAAATCCCTGCTACTAGGAGGTA 3914  
QY 3573 ATAATAAGGTTGAGTTTCCACAAAAAATAAAAAA 3611  
DB 3915 ATAATAAGGTTGAGTTTCCACAAAAAATAAAAAA 3953  
RESULT 2  
QY2520 ID Q92520 standard; cDNA to mRNA; 3962 BP.  
AC Q92520.  
DT 26-NOV-1995 (first entry)  
DE Human mammary carcinoma kinase 10 (MCK-10) cDNA.  
KW Mammary carcinoma kinase 10; transmembrane receptor;  
KW receptor tyrosine kinase; cancer; ss.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT Cds 321..3080 /\*tag= a  
FT misc\_difference 2315 /\*tag= b  
FT /\*note= "some clones have 6 AA deletion here"  
PN W09514088-A.  
PD 26-MAY-1995.  
PF 16-NOV-1994; E03797.  
PR 16-NOV-1993; US-153397.  
PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.  
PI Alves FHE, Ullrich A;  
DR WPI: 95-224054/29.  
DR P-PSDB; R75502.  
PT New nucleic acid encoding MCK-10 receptor tyrosine kinase - and  
PT derived vectors, transformed cells, proteins and antibodies useful  
PT for diagnosis and treatment of proliferative disease, esp. cancer,  
PT and for screening modulators  
PS Claim 4; Page 50-52; 115pp; English.  
CC cDNA prep'd. from human breast cancer cell line MCF7 (ATCC HTB22) and  
CC used in a PCR with two degenerate oligo primer pools based on  
CC conserved sequences of the kinase domain of receptor tyrosine  
CC kinases. One clone, designated MCK-10, was identified as novel RTK.  
CC The PCR fragment was used to screen a lambda gt11 library of human  
CC fetal brain cDNA. Several overlapping clones were identified. The  
CC composite of these cDNA clones is given in Q92520 and the deduced AA  
CC sequence in R75502. Some of the clones had a deletion of 6 AAs at  
CC position 2315 in the MCK-10 sequence. MCK-10 has all the  
CC characteristics of a receptor PTK (see R75502 FT). Screening of  
CC human placental library yielded two cDNA clones MCK-10-1 and  
CC MCK-10-2. One of the clones isolated from the human fetal brain  
CC library cont'd. an additional 18 nts in the TK domain. The MCK-10 splice  
CC isoforms have been designated MCK-10-1 (with an additional 111 bp between  
CC nts 1832 and 1943); MCK-10-2 (without any insertions); MCK-10-3 (with the  
CC additional 111 bp and 18 bp in the TK domain); and MCK-10-4 (with the  
CC additional 18 bp). The predicted mol. wts. of MCK-10-1 and MCK-10-2  
CC proteolipids are 101.13 and 97.17 kD respectively, and can thus be  
CC subdivided into a 34.31 kD alpha subunit and 66.84 or 62.88 kD  
CC beta subunits that contain the TK homology and alternative splice sites.  
CC Sequence 3962 BP; 735 A; 1235 C; 1181 G; 811 T;  
SQ  
Query Match 94.8%; Score 3449.4; DB 1; Length 3962;  
Best Local Similarity 97.0%; Pred. No. 0;  
Matches 3588; Conservative 0; Mismatches 6; Indels 105; Gaps 3;  
QY 17 GTTGACTTGAAGGAATGCCAAGAGATGCTGCCCCACCCCTTAGCCCGGAGGATCAG 76  
DB 256 GTTGACTTGAAGGAATGCCAAGAGATGCTGCCCCACCCCTTAGCCCGGAGGATCAG 315  
QY 77 GAGCTATGGGACAGAGGCGCTGTCTTACTGCTGTCTTGTGTGCAAGTGAG 136



QY 2279 CGCAGGGGCCACCATCAGTACCCATGCTGCTGATGTGGCAGCCAGATCGCTCG 2338  
 DB 2536 CGCAGGGGCCACCATCAGTACCCATGCTGCTGATGTGGCAGCCAGATCGCTCG 2595  
 QY 2339 GCATGGCTATCTGGCCACACTCAACTTGTACATCGGACCTGGCCACGCGAACTGCC 2398  
 DB 2596 GCATGGCTATCTGGCCACACTCAACTTGTACATCGGACCTGGCCACGCGAACTGCC 2655  
 QY 2399 TAGTTGGGAAATTTACCATCAAAATCGACATTTGGCATGAGCCGAACTCTATG 2458  
 DB 2656 TAGTTGGGAAATTTACCATCAAAATCGACATTTGGCATGAGCCGAACTCTATG 2715  
 QY 2459 CTGGGACTATTACCTGTGCGAGGCGCGGAGTGTGCTGCTGCTGCTGCTGCTGCTG 2518  
 DB 2716 CTGGGACTATTACCTGTGCGAGGCGCGGAGTGTGCTGCTGCTGCTGCTGCTGCTG 2775  
 QY 2519 AGTGCATCTCTATGGGGAAGTTTACGACCTCGAGTGCAGTGTGGGCTTTGGTGTGACCC 2578  
 DB 2776 AGTGCATCTCTATGGGGAAGTTTACGACCTCGAGTGCAGTGTGGGCTTTGGTGTGACCC 2835  
 QY 2579 TGTGGGAGTGTGATGCTCTGTAGGCGCCAGCCCTTTGGGAGTGTACCGACGAGCAGG 2638  
 DB 2836 TGTGGGAGTGTGATGCTCTGTAGGCGCCAGCCCTTTGGGAGTGTACCGACGAGCAGG 2895  
 QY 2639 TCATCGAGACGCGGGGAGTCTTCCGCGGACGAGCGCGGAGTGTACCTGTCCGGC 2698  
 DB 2896 TCATCGAGACGCGGGGAGTCTTCCGCGGACGAGCGCGGAGTGTACCTGTCCGGC 2955  
 QY 2699 CGCTGCTCTGCGCGAGGCGCTATATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 2758  
 DB 2956 CGCTGCTCTGCGCGAGGCGCTATATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 3015  
 QY 2759 AGCAGGACACCCCTTTCCAGTGCATCGGTTCCTGCGAGGAGTGTACCTCAACACGG 2818  
 DB 3016 AGCAGGACACCCCTTTCCAGTGCATCGGTTCCTGCGAGGAGTGTACCTCAACACGG 3075  
 QY 2819 TGTGAATCACATCAGTGCCTCTCCCTCAGGAGTGTACCGGAGGAGCAGTGTACA 2878  
 DB 3076 TGTGAATCACATCAGTGCCTCTCCCTCAGGAGTGTACCGGAGGAGCAGTGTACA 3135  
 QY 2879 CTAACAAAGAGACACAATGGCACCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2938  
 DB 3136 CTAACAAAGAGACACAATGGCACCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3195  
 QY 2939 AATAGAGCAGTGTGAGTGC----- 2959  
 DB 3196 AATAGAGCAGTGTGAGTGCAGGTGGGCTGGGCGCCACCGAGGAGTGTATGCCCTTCTC 3255  
 QY 2959 -----AGAGCCCTGTGCTG 2972  
 DB 3256 CCCTTCTTGACACACTCTATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3315  
 QY 2973 CCACCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3032  
 DB 3316 CCACCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3075  
 QY 3033 AAGGTGGGAGAAATATAGGATAGACACTGGACATGGCCCATTTGGAGCAGTGTGGCCCC 3092  
 DB 3376 AAGGTGGGAGAAATATAGGATAGACACTGGACATGGCCCATTTGGAGCAGTGTGGCCCC 3435  
 QY 3093 ACTGGACACACTGATCTCTGGAGAGTGGCTGCGCCCGCCAGCTTCTCTCTGCTGCTGCTG 3152  
 DB 3436 ACTGGACACACTGATCTCTGGAGAGTGGCTGCGC-----CCCGAGCTTCTCTCTGCTGCTGCTG 3494  
 QY 3153 ACTGGACCCCACTGGCTGAGAAATCTGGGGGTGAGGAGCAAGAGGAGGAGGAGGAGGAGG 3212  
 DB 3495 ACTGGACCCCACTGGCTGAGAAATCTGGGGGTGAGGAGCAAGAGGAGGAGGAGGAGGAGG 3554  
 QY 3213 TTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3272  
 DB 3555 TTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3614  
 QY 3273 GAAACACTGGAGCTGGGGGTAGCCCCCGCCCGCCAGCTCTAGTACACCCCTTCCGACCTG 3332

DB 3615 GAAACACTGGAGCTGGGGGTAGCCCCCGCCCGCCAGCTCTAGTACACCCCTTCCACTTG 3674  
 QY 3333 CAGTCTGTAGTGAAGTCTCTAAGCCTATAGCTTCTGTCTGTGGAGTAAATATGGGATT 3392  
 DB 3675 CAGTCTGTAGTGAAGTCTCTAAGCCTATAGCTTCTGTCTGTGGAGTAAATATGGGATT 3734  
 QY 3393 GGGGGAAAGAGGAGGAGCAAGCGCCATAGCTTGGGTTGGACATCTCTAGTGTAGCTGC 3452  
 DB 3735 GGGGGAAAGAGGAGGAGCAAGCGCCATAGCTTGGGTTGGACATCTCTAGTGTAGCTGC 3794  
 QY 3453 CACATTGATTTTCTAATCACTTGGGTTGTGTACATTTTGGGGGAGAGACACAGAT 3512  
 DB 3795 CACATTGATTTTCTAATCACTTGGGTTGTGTACATTTTGGGGGAGAGACACAGAT 3854  
 QY 3513 TTTTACACTAATATATGACCTAGCTTGGGCAATTTTATCCCTGCTAGTGTAGCTGTA 3572  
 DB 3855 TTTTACACTAATATATGACCTAGCTTGGGCAATTTTATCCCTGCTAGTGTAGCTGTA 3914  
 QY 3573 ATAATAAGTTGAGTTTTCACAAAAA 3611  
 DB 3915 ATAATAAGTTGAGTTTTCACAAAAA 3953

## RESULT 3

Q92522 standard; cDNA to mRNA; 3962 BP.

AC Q92522;

DE 26-NOV-1995 (first entry)

DE Human mammary carcinoma kinase 10 (MCK-10) cDNA.

KW Mammary carcinoma kinase 10; MCK-10; transmembrane receptor;

OS receptor tyrosine kinase; cancer; ss.

OS Homo sapiens.

FH Key Location/Qualifiers

FT cds 321..3080

FT misc\_difference 2315 /tag= a

FT /tag= b

FT /note= "some clones have AA deletion here"

PN WO9514089-A.

PD 26-MAY-1995.

PF 16-NOV-1994; E03799.

PR 16-NOV-1993; US-153397.

PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.

PI Alves FHE, Ullrich A;

DR WPI: 95-224055/29.

DR P-PSDB: R75504.

PT New nucleic acid encoding CK-2 receptor tyrosine kinase - and  
 PT derived vectors, transformed cells, proteins and antibodies, useful  
 PT for diagnosis and treatment of proliferative and nervous system  
 PT diseases and for screening modulators  
 PS Disclosure: Page 67-69; 115pp; English.

CC cDNA prep. from human breast cancer cell line MCF7 (ATCC HTB22) was  
 CC used in a PCR with two degenerate oligo primer pools based on  
 CC conserved sequences of the kinase domain of receptor tyrosine  
 CC kinases. One clone, designated MCK-10, was identified as novel RTK.  
 CC The PCR fragment was used to screen a lambda gt11 library of human  
 CC fetal brain cDNA. Several overlapping clones were identified. The  
 CC composite of these cDNA clones is given in Q92522 and the deduced AA  
 CC sequence in R75504. Some of the clones had a deletion of 6AA at posn.  
 CC 2315 in the MCK-10 sequence. MCK-10 has all the characteristics of  
 CC a receptor PTK (see R75504 FT). Screening of human placental library  
 CC yielded two cDNA clones. One of the clones isolated from the human  
 CC fetal brain library contained an additional 18 nts in the TK  
 CC domain. The MCK-10 splice isoforms have been designated MCK-10-1  
 CC (with an additional 111 bp between nts 1832 and 1943); MCK-10-2  
 CC (without any insertions); MCK-10-3 (with the additional 111 bps and  
 CC 18 bp in the TK domain); and MCK-10-4 (with the additional 18 bp).  
 CC The predicted mol. wts. of MCK-10-1 and MCK-10-2 proteolipids are  
 CC 101.13 and 97.17 kD respectively, and can thus be subdivided into a  
 CC 34.31 kD alpha subunit and a 66.84 or 62.88 kD beta subunits that  
 CC contain the TK homology and alternative splice sites.  
 CC Sequence 3962 BP; 735 A; 1235 C; 1181 G; 811 T;

SQ

Query Match 94.8%; Score 3449.4; DB 1; Length 3962;  
Best Local Similarity 97.0%; Pred. No. 0;  
Matches 3588; Conservative 0; Mismatches 6; Indels 105; Gaps 3;

QY	17	GTGGACATTGAAGGAATGCCAAGAGATGCTGCCCCACCCCTTAGGCCGAGGATACG	76	1037	GCCACAACCTAGGGGCAACCTGGGGAGCCCAAGAGCCCGGGCTGTCTCACTGCCCCCTTG	1096
DB	256	GTGGACATTGAAGGAATGCCAAGAGATGCTGCCCCACCCCTTAGGCCGAGGATACG	315	1276	GCCACAACCTAGGGGCAACCTGGGGAGCCCAAGAGCCCGGGCTGTCTCACTGCCCCCTTG	1335
QY	77	GAGCTATGGGACCAGAGGCCCTGTCACTTTACTGCTGCTGCTCTTGGTGCAAGTGGAG	136	1097	GGGGCGGTGTGCTGCTGCTTTCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1156
DB	316	GAGCTATGGGACCAGAGGCCCTGTCACTTTACTGCTGCTGCTCTTGGTGCAAGTGGAG	375	1336	GGGGCGGTGTGCTGCTGCTTTCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1395
QY	137	ATGCTGACATGAAGGACATTTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	196	1157	TCAGGGAATCT	1216
DB	376	ATGCTGACATGAAGGACATTTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	435	1396	TCAGGGAATCT	1455
QY	197	ACCGGACATCCAGACAGTACATCTCTGCTTCCAGTCTCTGCTGCTGCTGCTGCTGCTGCT	256	1217	CCTTCCCGGACAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1276
DB	436	ACCGGACATCCAGACAGTACATCTCTGCTTCCAGTCTCTGCTGCTGCTGCTGCTGCTGCT	495	1456	CCTTCCCGGACAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1515
QY	257	CCCGCCACAGAGTTGGAGAGCAGTACGCGGATGGGGCTGCTGCTGCTGCTGCTGCTGCTGCT	316	1277	AGCTGGAGCCAGAGCCAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1336
DB	496	CCCGCCACAGAGTTGGAGAGCAGTACGCGGATGGGGCTGCTGCTGCTGCTGCTGCTGCTGCT	555	1516	AGCTGGAGCCAGAGCCAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1575
QY	317	TGTTTCCCAAGGAGGAGTACTTGCAGTGGATCTACACGACTACACTGCTGCTGCTGCTGCT	376	1337	TCATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1396
DB	556	TGTTTCCCAAGGAGGAGTACTTGCAGTGGATCTACACGACTACACTGCTGCTGCTGCTGCT	615	1576	TCATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1635
QY	377	TGGTGGGACCCAGGAGCGATCCGGGGGCTGGGCAAGAGTCTTCCCGAGCTAC	436	1397	GGCGGTGCATCTGGGCGAGGCTCTCAGCAAGGCTGAACGAGGCTGTTGGAAGAGGAGC	1456
DB	616	TGGTGGGACCCAGGAGCGATCCGGGGGCTGGGCAAGAGTCTTCCCGAGCTAC	675	1636	GGCGGTGCATCTGGGCGAGGCTCTCAGCAAGGCTGAACGAGGCTGTTGGAAGAGGAGC	1695
QY	437	GGCTCGGTTACTCCCGGATGGTGGCCCTGGATGGGCTGGAAGGACCGCTGGGGTCAGG	496	1457	TGACGGTTCACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1516
DB	676	GGCTCGGTTACTCCCGGATGGTGGCCCTGGATGGGCTGGAAGGACCGCTGGGGTCAGG	735	1696	TGACGGTTCACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1755
QY	497	AGGTGATCTCAGGCAATGAGGACCTGAGGAGTGGTCTGAAGGACCTTGGGGCCCCCA	556	1517	GAGAGCCACCCCGCTACAGAGCCCGGCTGCTGGAATCCGCGCCACTCCGCTCCCT	1576
DB	736	AGGTGATCTCAGGCAATGAGGACCTGAGGAGTGGTCTGAAGGACCTTGGGGCCCCCA	795	1756	GAGAGCCACCCCGCTACAGAGCCCGGCTGCTGGAATCCGCGCCACTCCGCTCCCT	1815
QY	557	TGGTGGCGGACTGCTGCTTCTACCCCGGGCTGACCGGCTATGAGCGTCTGCTGCTGCTGCT	616	1577	GTGTCCCAATGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1636
DB	796	TGGTGGCGGACTGCTGCTTCTACCCCGGGCTGACCGGCTATGAGTCTGCTGCTGCTGCTGCT	855	1816	GTGTCCCAATGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1875
QY	617	GGGTAGAGCTCTATGGCTGCTCTGGAGGATGGACTCTCTCTTACACGCCCTCTGCTGCTGCT	676	1637	CTTAGCGGCTCCCTCGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1696
DB	856	GGGTAGAGCTCTATGGCTGCTCTGGAGGATGGACTCTCTCTTACACGCCCTCTGCTGCTGCT	915	1876	CTTAGCGGCTCCCTCGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1935
QY	677	GGCAGACAATGATTTATCTGAGGCGGTGTACCTCAAGCACTCCACCTATGACGACATA	736	1697	ACACCCAGGCTACAGTGGGAGCTATATGAGGCTGAGAAGCCAGGCGCGGCTCTGCTGCTGCT	1756
DB	916	GGCAGACAATGATTTATCTGAGGCGGTGTACCTCAAGCACTCCACCTATGACGACATA	975	1936	ACACCCAGGCTACAGTGGGAGCTATATGAGGCTGAGAAGCCAGGCGCGGCTCTGCTGCTGCT	1995
QY	737	CCGTGGGCGGACTGACGATATGGGGTCTGGGCGAGCTGGGAGATGGTGGGGGCTGG	796	1757	CGCCAGCTCCCAAGAACAGCGTCCCGCATTTATGCGAGGCTGACATTTGTACCCCTGCAGG	1816
DB	976	CCGTGGGCGGACTGACGATATGGGGTCTGGGCGAGCTGGGAGATGGTGGGGGCTGG	1035	1996	CGCCAGCTCCCAAGAACAGCGTCCCGCATTTATGCGAGGCTGACATTTGTACCCCTGCAGG	2055
QY	797	ATGACTTTAGGAAGAGTCAGGAGCTGCGGGTCTGGCCAGGCTATGACTATGTGGGTGGA	856	1876	GGGTACCCGGGCGCAACCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1876
DB	1036	ATGACTTTAGGAAGAGTCAGGAGCTGCGGGTCTGGCCAGGCTATGACTATGTGGGTGGA	1095	2056	GGGTACCCGGGCGCAACCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2115
QY	857	GCAACACAGCTTCTCCAGTGGCTATGTTGAGATGGAGTGTGAGTTTACCGGCTGAGG	916	1877	GGCCCCCAGAGTGGATTTTCCCTCGATCTGACTCGGCTTCAAGGAGAGCTTGGCGAGG	1936
DB	1096	GCAACACAGCTTCTCCAGTGGCTATGTTGAGATGGAGTGTGAGTTTACCGGCTGAGG	1155	2116	GGCCCCCAGAGTGGATTTTCCCTCGATCTGACTCGGCTTCAAGGAGAGCTTGGCGAGG	2175
QY	917	CTTTCAGGCTATGAGGTCCACTGTACACATGCACACCTGGGAGCCGCTGCTGCTGCTGCTGCT	976	1937	GGCAGTTTGGGAGGTGCACCTGTGTGAGGTGCAGCGCTCAAGATCTGCTGCTGCTGCTGCTGCTGCT	1996
DB	1156	CTTTCAGGCTATGAGGTCCACTGTACACATGCACACCTGGGAGCCGCTGCTGCTGCTGCTGCT	1215	2176	GGCAGTTTGGGAGGTGCACCTGTGTGAGGTGCAGCGCTCAACATCTGCTGCTGCTGCTGCTGCTGCT	2235
QY	977	GGGGGTGGAATGTCGCTTCCGGCTGGCCCTGCCATGGCTGGGAGGGGAGCCCATGC	1036	1997	ATTTCCCTTAAATGTCGCTAAGGGACACCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2056
DB	1216	GGGGGTGGAATGTCGCTTCCGGCTGGCCCTGCCATGGCTGGGAGGGGAGCCCATGC	1275	2236	ATTTCCCTTAAATGTCGCTAAGGGACACCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2295
				2057	CAGATGCCACCAAGATG-----CCAGGAATGATTTCTTGAAGAGG	2098
				2296	CAGATGCCACCAAGATGCTCTCTCTGCTTCTTCCAGGAATGATTTCTTCTGAAAGAGG	2355

Qy	2099	TGAAGATCATGTGCGAGCCTCAAGGACCCAAACATCATTTCCGGCTGCTGGCGCTGTGTGTGC	2158
Db	2356	TGAAGATCATGTGCGAGCCTCAAGGACCCCAACATCATTTCCGGCTGCTGGCGCTGTGTGTGC	2415
Qy	2159	AGAGACACCCCTCTCATGATTACTGTACTACATGAGAACGGCGAGCTCAACAGTTCC	2218
Db	2416	AGAGACACCCCTCTCATGATTACTGTACTACATGAGAACGGCGAGCTCAACAGTTCC	2475
Qy	2219	TCAGTGCACACAGCTGSGAGGACAAAGGACGCGAGGGGGCCCTTGGGACGGCAGCGTG	2278
Db	2476	TCAGTGCACACAGCTGSGAGGACAAAGGACGCGAGGGGGCCCTTGGGACGGCAGCGTG	2535
Qy	2279	CGAGGGGGCCACCATCAGCTACCCAAATGCTCTGCTGCATGTGGGAGCCAGATCGCCTCG	2338
Db	2536	CGAGGGGGCCACCATCAGCTACCCAAATGCTCTGCTGCATGTGGGAGCCAGATCGCCTCG	2595
Qy	2339	GCATGGCTATCTGGCCACACCTCAACTTTGTACATCGGAGCTTGGCACGCGGAACCTGCC	2398
Db	2596	GCATGGCTATCTGGCCACACCTCAACTTTGTACATCGGAGCTTGGCACGCGGAACCTGCC	2655
Qy	2399	TAGTTGGGAAAAATTTCCACCATCAAAATCGACACTTTGGCATGAGCGGAAACCTCTATG	2458
Db	2656	TAGTTGGGAAAAATTTCCACCATCAAAATCGACACTTTGGCATGAGCGGAAACCTCTATG	2715
Qy	2459	CTGGGACATATTACCTGTGCGAGGCGCGGACGTGCTGCCAATCCGCTGGATGBCCTGGG	2518
Db	2716	CTGGGACATATTACCTGTGCGAGGCGCGGACGTGCTGCCAATCCGCTGGATGBCCTGGG	2775
Qy	2519	AGTGCACTCTCATGGGAAAGTTACAGACTGCGAGTGACGTGTGGCCCTTTGGTGTGACC	2578
Db	2776	AGTGCACTCTCATGGGAAAGTTACAGACTGCGAGTGACGTGTGGCCCTTTGGTGTGACC	2835
Qy	2579	TGTGGGAGTGTGATGCTCTCTAGGGCCACGCCCTTTGGGACGTCAACGACGACGAGG	2638
Db	2836	TGTGGGAGTGTGATGCTCTCTAGGGCCACGCCCTTTGGGACGTCAACGACGACGAGG	2895
Qy	2639	TCATCAGAAACGGGGGAGTTCTTCGGGACACAGGGCCGCGAGGTACCTGTCCCGGC	2698
Db	2896	TCATCAGAAACGGGGGAGTTCTTCGGGACACAGGGCCGCGAGGTACCTGTCCCGGC	2955
Qy	2699	CGCTGCCTGCCGCGAGGCCCTATATGAGCTGATGCTTTCGGCTGGAGCGGGAGTCTG	2758
Db	2956	CGCTGCCTGCCGCGAGGCCCTATATGAGCTGATGCTTTCGGCTGGAGCGGGAGTCTG	3015
Qy	2759	AGCAGGACACACCTTTTCCAGCTGCATCGGTTCTTCCGACAGAGATGCACTCAACACGG	2818
Db	3016	AGCAGGACACACCTTTTCCAGCTGCATCGGTTCTTCCGACAGAGATGCACTCAACACGG	3075
Qy	2819	TGTGAATCACATTCAGCTGCCCTCCCTCAGGAGTGTATCCAGGGAAGCCAGTGACA	2878
Db	3076	TGTGAATCACATTCAGCTGCCCTCCCTCAGGAGTGTATCCAGGGAAGCCAGTGACA	3135
Qy	2879	CTAAACAAGAGACACAATGGCACTCTGCCCTTCCCTCCGACAGCCCATCACCTCT	2938
Db	3136	CTAAACAAGAGACACAATGGCACTCTGCCCTTCCCTCCGACAGCCCATCACCTCT	3195
Qy	2939	AATAGGCGAGTGAGACTGC-----	2959
Db	3196	AATAGGCGAGTGAGACTGC-----	3255
Qy	2959	-----	2972
Db	3256	CCCTTCTCGACACACTCTATGTCCTTCCCTTCTGTTCTTCTCTAGAGCCCTTGTG	3315
Qy	2973	CCACCCAGCTGTCTGTGGATGGATCCTCTCCACCTCTCTAGCCATCCCTTGGG	3032
Db	3316	CCACCCAGCTGTCTGTGGATGGATCCTCTCCACCTCTCTAGCCATCCCTTGGG	3375
Qy	3033	AAGGTTGGGAGAAATATAGGATAGACCTGGACATGCCCATTTGGAGCACTTGGGCCCC	3092
Db	3376	AAGGTTGGGAGAAATATAGGATAGACCTGGACATGCCCATTTGGAGCACTTGGGCCCC	3435
Qy	3093	ACTGGACAACATGTATCTCGAGAGGTGGCTGCGCCCCCAGCTTCTCTCCCTGTCA	3152

3436	Db	ACTGGCAAACTGATCTCTGGAGAGGTGGCTGG-CCCAGCTTCTCTCTCCCTGTCCAC	3494
3153	QY	ACATGGACCCCACTGGCTGAGAACTCTGGGGTGAGGAGGACAGAAGGAGGAGAAATG	3212
3495	Db	ACATGGACCCCACTGGCTGAGAACTCTGGGGTGAGGAGGACAGAAGGAGGAGGAAATG	3554
3213	QY	TTTCCTTGTGGCTGCTCTGTACTTGTCTCAGCTTGGGCTTCTTCCTCTCCCACTCACT	3272
3555	Db	TTTCCTTGTGGCTGCTCTGTACTTGTCTCAGCTTGGGCTTCTTCCTCTCCCACTCACT	3614
3273	QY	GAACACATGACCTGGGGGTAGCCCGCCAGCCCTCAGTCAACCCCACTTCCCACTG	3332
3615	Db	GAACACATGACCTGGGGGTAGCCCGCCAGCCCTCAGTCAACCCCACTTCCCACTG	3674
3333	QY	CAGCTCTGTAGCTAGAACTTCTTAAGCCCTATACGTTTCTGTGGAGTAAATATTGGGATT	3392
3675	Db	CAGCTCTGTAGCTAGAACTTCTTAAGCCCTATACGTTTCTGTGGAGTAAATATTGGGATT	3734
3393	QY	GGGGGAAAGAGGAGGACAAAGCCCATAGCCCTGGGGTTGGACATCTCTAGTGTAGCTGC	3452
3735	Db	GGGGGAAAGAGGAGGACAAAGCCCATAGCCCTGGGGTTGGACATCTCTAGTGTAGCTGC	3794
3453	QY	CACATTGATTTTCTATATCACTCTGGGGTTTGTACATTTTGGGGGAGAGACACAGAT	3512
3795	Db	CACATTGATTTTCTATATCACTCTGGGGTTTGTACATTTTGGGGGAGAGACACAGAT	3854
3513	QY	TTTTTACACTAATATATGGACCTAGCTTTGAGGCAATTTTAAATCCCTGCACATAGGCAGTA	3572
3855	Db	TTTTTACACTAATATATGGACCTAGCTTTGAGGCAATTTTAAATCCCTGCACATAGGCAGTA	3914
3573	QY	ATAATAAGGTTGAGTTTCCACAAAAAATAAAAAA	3611
3915	Db	ATAATAAGGTTGAGTTTCCACAAAAAATAAAAAA	3953

## RESULT.T

AC	Q84782	standard; DNA; 3754 BP.
AC	Q84782;	
DE	17-AUG-1995	(first entry)
DE	Protein-tyrosine-kinase PTK22.	
KW	Protein-tyrosine-kinase; PTK; discoidin domain receptor; cancer;	
KW	breast tumor; mamma carcinoma; diagnosis; prognosis; therapy; ds.	
OS	Homo sapiens.	
FT	Key	Location/Qualifiers
FT	cds	142..2886
FT		/*tag= a
PN	W09502187-A.	
PD	19-JAN-1995.	
PF	08-JUL-1994; G01480.	
PR	09-JUL-1993; GB-014271.	
PA	(WELL-) CANCER RES INST.	
PA	(CANC-) WELLCOME FOUND LTD.	
PI	Barker KT, Crompton MR, Gusterson BA, Martindale JE;	
PI	Mitchell PJ, Page MJ, Spence P;	
DR	WPI: 95-066991/09.	
DR	P-PSDB: R71100.	
PT	Method for screening substances, using protein tyrosine kinase -	
PT	for potential utility as therapeutic agents for cancer	
PS	Claim 1; Page 26-30; 51pp; English.	
CC	cDNA derived from tumor metastatic tissue was amplified using	
CC	primers (given in Q84783-84) based on sequences (R71101, R71103)	
CC	associated with protein-tyrosine-kinases (PTK). Novel PTK22 was	
CC	identified in an isolated subclone. The 3' sequence of PTK22 was	
CC	obtained by reverse transcription (using the primer of Q84786) and	
CC	PCR amplification (primers Q84787-88) of RNA of human breast	
CC	carcinoma cell line MDA MG 468. The partial DNA sequence of PTK22	
CC	is given in Q84782.	
SC	Sequence 3754 BP;	713 A; 1145 C; 1121 G; 775 T;

Query Match 94.4%; Score 3433.8; DB 1; Length 3754;



Query Match 18.4%; Score 669.4; DB 1; Length 3120;  
Best Local Similarity 56.9%; Pred.No.1.5e-127;  
Matches 153; Conservative 0; Mismatches 192; Gaps 10;

QY 109 CTGCTGCTGCTCTTGGTGCGAAGTGGAGATGCTGCACATGAAGGACATTTCATCTGCC 168  
|||||  
db 509 CTGTGCTGCTCCTCTTTGCTCATCTCTGGTGTCTCAAAAGCTCAGGTTATTCAGCC 568







Db	2549	ATGGCAACCCAGATTCCCTCTGGTATGAAGTACCTTTTCGTCTCAACTTTGTCCACCGA	2608
Qy	2377	GACCTGGCCACCGGAACCTCCCTAGTTGGGAAAAATTTACCACTCAAAATCGCAGACTTT	2436
Db	2609	GATCTGCCACACAGAACTGTTTGTAGTGGCAAGAATTACACCAATCAAGATAGCTGATTTT	2668
Qy	2437	GGCATGAGCCGGAACCTCTATCCTGGGACATATACCCTGTGTGAGGGCCCGGGCAGTGCCTG	2496
Db	2669	GGCATGAGCAGAAACCTGTACAGTGGTGATTACTACCGGATCCAGAGCCCGGGCGGTGCTC	2728
Qy	2497	CCCATCCGTGGATGSCCTGGGAGTGCATCTCATGTGGGAAAGTTTCAGGATGCGAGTGCAC	2556
Db	2729	CCCATTCGCTGGATGTCCTGGAAAGCATCTTGTGTGGCAAAATTCACCGGCAAGTGAT	2788
Qy	2557	GTCTGGCCCTTTGGTGTGACCCCTGTGGAGGTGCTGATGCTCTGTAGGGCCCGACCCCTTT	2616
Db	2789	GTGTGGCCCTTTGGGTGTACTGTGTGGAGACCTTTCACCTTTTCCAGAGACAGCCCAT	2848
Qy	2617	GGCAGCTCACCAGCAGCAGCAGTCTATCGAGAACGCGGGGAGTTCTTCCGGGACACAGGC	2676
Db	2849	TCCAGCTGTCSGATGAGCAGGTTATCGAGAACACTGGAGAGTTCTTCCGAGACCAAGGG	2908
Qy	2677	CGCAGGTGTACTGTCCCGGCCCTCGCTTCCCGCAGGSCCTATATGAGCTGATGCTT	2736
Db	2909	AGCAGATCTATCTCCCTCAACACAGCCCTTTTGCOCGACTGTGTGTATAAGCTGATGCTC	2968
Qy	2737	CGGTGCTGGAGCCGGGAGTCTGAGCAGCGACACCCCTTTTCCCAAGCTGCATCGTGTCCCTG	2796
Db	2969	AGCTGCTGGAGAGAGAACCAAGCACCGGCCATCTTCCAGGAAATACACCTTCCTGCTT	3028
Qy	2797	GCAGAGGTGCTACTCAACACGGTGTGAATCA	2827
Db	3029	CTTCAACAGAGCCGAGTGATGATGATCA	3059

## RESULT 6

```

V65317
ID V65317 standard; DNA; 3120 BP.
AC V65317;
DE 22-JAN-1999 (first entry)
DE Receptor protein tyrosine kinase (PTK) subtype tyro-10 encoding DNA.
KW PTK; receptor; protein tyrosine kinase; recombinant; grafting;
KW diagnosis; tumour; skin transplant; connective tissue; tyro-10; ss.
OS Rattus sp.
FT Key Location/Qualifiers
FT 485..3049
FT CDS /*tag= a
      /product= "PTK subtype tyro-10"
      US5937448-A.
PD 17-NOV-1998.
PD 02-MAY-1994; 237401.
PR 15-MAY-1992; US-884486.
PR 02-MAY-1994; US-237401.
PA (SALK ) SALK INST BIOLOGICAL STUDIES.
PI Lai CHC, Lemke GE;
DR WPI; 99-023436/02;
DR P-P8DB; W81409.
DR PT Nucleic acids encoding protein tyrosine kinase sub:types - for
PT identification of new sub:types and treatment of diseases associated
PT with the kinase
PS Claim 1: Columns 53-58; 47pp; English.
CC This DNA encodes a receptor protein tyrosine kinase (PTK) subtype
CC tyro-10. The invention provides sequences V65308 to V65313, V65315, and
CC V65317 to V65319 that encode proteins having a tyrosine kinase domain and
CC a tissue expression pattern of a receptor PTK subtype selected from
CC tyro-1, tyro-2, tyro-3, tyro-4, tyro-5, tyro-6, tyro-8, tyro-10, tyro-11,
CC and tyro-12, respectively. The polynucleotides are useful for the
CC detection of tyrosine kinase domain sequences and detection of tissue
CC expression patterns of PTK subtypes. The cDNAs can also be injected into
CC oocytes, the protein expressed, and expression products screened for
CC using antibodies against tyrosine kinase epitopes. These subtypes
CC sequences can be used for the design of oligonucleotides, for use in
CC amplification reactions to isolate other subtype sequences. These

```

QY 1000 CGTGGCCCTGCTGCTGGAGGGGAGCCCATCGGCCAACAACTAGGGGGCAACCTG 1059  
 Db 1406 TCGG---AAGCCAGGAGTGGGAACCACTGCTGTCTACTTTCCCTGCTGCTGACGAT 1462  
 QY 1060 GGGAGCCAGAGCCGGGCTGCTCAGTGGCCCTTGGGGCGGCTGCTGCTGCTGCTG 1119  
 Db 1463 GTGAACCCAGTCCCGGGTGTGACGGGTGGCCCTCCACCCGAATGGCCAGTGGCCATC 1522  
 QY 1120 CAGTGGCGCTCTCTTTGGGGGGCTGCTTACTCTTCCAGGAAATCTCTTCTATCT 1179  
 Db 1523 AGTGCCAAATACCAATTTGGCCAGACAGTGGATGATGTTTCCAGGAGATCACTTCCAA 1582  
 QY 1180 GATGTGGTGAACAATCTCTCTCGGACCTGGGAGGACCTTCCCGCCAGCCCTGGTGG 1239  
 Db 1583 GATGTGCAATGATAACAACCTGAGAGCCCTTCCCACTCTCTA----- 1629  
 QY 1240 CCGCTGGCCCACTCCCACTTTCAGCAGCTTGGAGCTGGAGCCAGAGCCAGCAG 1299  
 Db 1629 -----TGGCACCCACCACTATGAT 1648  
 QY 1300 CCGTGGCCCAAGCCGAGGGAGCCGACCGCATCTCTGCTGCTGCTGCTGCTGCTG 1359  
 Db 1649 CCATGCTTAAAGTGTATGATACCACTCGGATCCTGATGCTGCTGCTGCTGCTG 1708  
 QY 1360 ATCTGCTCTGCTGCTATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1419  
 Db 1709 ATCTTCTCTGCTGCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1768  
 QY 1420 CTGAGCAAGCTGAAGGGGTTGGAGAGAGCTGACGCTTCCACCTCTGCTGCTG 1479  
 Db 1769 CTAGAAAGGCTTACGAGAGGATGCTGGATGATGAATGACATGACCTTCCCTG 1828  
 QY 1480 GGGGACACTTCTCTCAACAACCGCCAGGCTCTAGAGAGCCACCCCGCTACAGAG 1539  
 Db 1829 AGCAGTCCAGATGTTCAATACACCGCTCTCT-----ATCACCAAGTGAACAGAG 1882  
 QY 1540 CCGGGGCTCGTGGGAATCGCCCCACTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1599  
 Db 1883 TCAACTCTACTTATGATGAATCTTCCCTTCCGCTGCTGCTGCTGCTGCTGCTG 1942  
 QY 1600 CTGCTCTCAATCCAGCTACCGCTCTCTTCTGCTGCTGCTGCTGCTGCTGCTG 1659  
 Db 1943 CTGATC----- 1949  
 QY 1660 CCGGGCCCCCACCACCGCTGGGCAACCCCAACACCCAGGCTTACAGTGGGAC 1719  
 Db 1949 -----CGAAGCTTCCAGAGTTTCTCCAGGAGAGGAGGA 1983  
 QY 1720 TATATGAGCGCTGAGAGCGGCGGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1779  
 Db 1984 GTCAAGGTGAGTGGTGTGTTGTAAGCGGGC-----CAGCCCAATGGACCTGAGGCGTG 2038  
 QY 1780 CCCCATTATCGAGGCTGACATGTTTACCTGCGAGGGCTCACGGGGGCAACACCTAT 1839  
 Db 2039 CCCCATTATCGAGGCTGAGAGCGGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTG 2098  
 QY 1840 GCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1896  
 Db 2099 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2158  
 QY 1897 CTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1956  
 Db 2159 CCGAGGAACTGTTGGGCTTCAAGGAGAGCTGCGGAGAGGCGGCTTGGGGAGTTCAT 2218  
 QY 1957 CTGTGTGAGTGCAGACCGCTTCAAGATCTGCTGCTGCTGCTGCTGCTGCTGCTG 2016  
 Db 2219 CTGTGTGAGTGCAGGAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2278  
 QY 2017 AAGGAGACCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2076  
 Db 2279 GCGAACCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2338  
 QY 2077 AGGAATGATTCTGAAAGAGGTGAAGATCATGCTGAGGCTCAAGGACCCCAACATCAT 2136

Db 2339 AGGAATGATTCTTAAAGGAGATCAAGATCATGCTCGGCTCAAGGACCCAAACATCATC 2398  
 QY 2137 CGGCTGCTGGCGTGTGTGAGGAGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2196  
 Db 2399 CGTCTTGTAGTGTGTGCTGCTGAGGACCCGCTGCTGCTGCTGCTGCTGCTGCTG 2458  
 QY 2197 AACGGCAGCTCAACCACTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2256  
 Db 2459 AATGAGATCTTAACTAGTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2502  
 QY 2257 GCCCTGGGAGGAGGAGGCTGCGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2316  
 Db 2502 -----GTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2548  
 QY 2317 GTGCAGCCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2376  
 Db 2549 ATGCAACCCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2608  
 QY 2377 GACCTGCCAGCGGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2436  
 Db 2609 GATCTGGCCACACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2668  
 QY 2437 GGCATGAGCGGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2496  
 Db 2669 GGCATGAGCAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2728  
 QY 2497 CCCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2556  
 Db 2729 CCCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2788  
 QY 2557 GTGTGGCGCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2616  
 Db 2789 GTGTGGCGCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2848  
 QY 2617 GGGCAGCTCACCGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2676  
 Db 2849 TCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2908  
 QY 2677 CGGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2736  
 Db 2909 AGGAGATCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2968  
 QY 2737 CGGTGCTGAGCGGAGTGTGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2796  
 Db 2969 AGTGTGGAAGAGAAACCAAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3028  
 QY 2797 GCAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2827  
 Db 3029 CTTGAGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3059

RESULT 7

Q92521 standard; cDNA to mRNA; 3157 BP.  
 ID Q92521.  
 AC Q92521.  
 DE 26-NOV-1995 (first entry)  
 DE Human colonic adenocarcinoma kinase 2 (CKK-2) cDNA.  
 KW Mammary carcinoma kinase 10; MCK-10; transmembrane receptor; CKK-2;  
 KW receptor tyrosine kinase; colonic adenocarcinoma kinase 2; cancer; ss.  
 OS Homo sapiens.  
 FH key Location/Qualifiers  
 FT cds 370..2937  
 FT /\*tag= a  
 PN W09514088-A.  
 PD 26-MAY-1995.  
 PF 16-NOV-1994; E03797.  
 PR 16-NOV-1993; US-153397.  
 PA (PLAC) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.  
 PI Alves FHE, Ullrich A;  
 DR WPI; 95-224054/29.  
 DR P-PSDB; R75503.  
 PT New nucleic acid encoding MCK-10 receptor tyrosine kinase - and

PT derived vectors, transformed cells, proteins and antibodies useful  
PT for diagnosis and treatment of proliferative disease, esp. cancer,  
PT and for screening modulators

PS Disclosure: Page 56-57; 115pp; English.  
CC A member of the mammary carcinoma kinase 10 (MCK-10) receptor  
CC tyrosine kinase family was identified using a PCR (with two degenerate  
CC oligo primer pools based on conserved sequences of the kinase domain of  
CC receptor tyrosine kinases) and cDNA prep. from colonic  
CC adenocarcinoma RNA. The nt sequence of the novel receptor,  
CC designated CKK-2, is given in Q92521 and the deduced AA sequence in  
CC R75503. Analysis of the CKK-2 nt and AA sequence indicated  
CC significant homology with MCK-10 throughout the extracellular,  
CC transmembrane and intracellular regions. The regions of homology  
CC extend into the N-terminus consensus sequence for the discoidin I  
CC like family of proteins.

SQ Sequence 3157 BP: 790 A; 802 C; 759 G; 806 T;

Query Match 18.3%; Score 666.8; DB 1; Length 3157;  
Best Local Similarity 56.8%; Pred. No. 5e-127;  
Matches 1532; Conservative 0; Mismatches 977; Indels 189; Gaps 9;

QY 110 TGTGCTGCTCTGTGGCAAGTGGAGATGCTGACATGAAGGACATTTTGTATCTCTGCCA 169  
DB 395 TGTGCTGCTCTGTGGCAAGTGGAGATGCTGACATGAAGGACATTTTGTATCTCTGCCA 454  
QY 170 AGTGGCGCTATGCGCTGGCATGACAGACCGGACCATCCAGACAGTGCATCTCTGCTT 229  
DB 455 TATGGCGCTATCTCTGGCATGTCAGGAGCCAGATTCCAGATGAGGACATCAGCTT 514  
QY 230 CAGCTCTCTGTCAGATTCCTGCGGCGCCGACAGCAGGTTGGAGAGCAGTGCAGGG 289  
DB 515 CCAGTCAAGTGGTCCAGAGTCCAGAGTGCACAAATATGGAAGGCTGGACTCAGAAGAAGGG 574  
QY 290 ATGGCGCTGGTGGCGGCGGAGGCTGCTTCCCAAGA---GGAGGAGTACTTGCAGG 346  
DB 575 ATGGAGCTTGGTGGCGGCGGAGGCTGCTTCCCAAGA---GGAGGAGTACTTGCAGG 634  
QY 347 TGGATCTACAAGCACTGCACCTGGTGGCTCTGTGGGACACCGGAGGAGGATGCGGGG 406  
DB 635 TTGACTTGCACACCTTCCATTTATCAGTCTGTGGGAGCCAGGGGCGCATGAGGAG 694  
QY 407 GCTGGGCAAGAGTCTCCGGAGCTACCGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGT 466  
DB 695 GTATGGCATCGAGTTGGGCGGCGGATGCAAGATCAATTTACAGTGGGATGTCAGTCT 754  
QY 467 GATGGCTGGAGGAGCCCTGGGTCAGGAGTGTCTCAGGCAATGAGGACCTTGAGG 526  
DB 755 GATCTCTGGGGAACCCCTGATGGGAACAGAGTGTGGATGGAATAGTAACCCCTATG 814  
QY 527 GATGGTGTGAAGGACCTTGGGCGGCGGCGGATGTTGGGCGGCTGCTTCTACCCCG 586  
DB 815 ACATTTTCTAAAGGACTTGGAGCGCCCATTTAGCCAGATTTGCGGTTTCTATCCAG 874  
QY 587 GGCTGACCGGGTATGACGCTGTCTCGGGGTAGAGTCTATGGCTGCTCTGGAGG 646  
DB 875 TCACCGCACTCCATGATGTGTATGAGAGTGGAGCTTTACGCTGTGTCTGGCTAG 934  
QY 647 ATGGACTCTGTCTTACACCGCCCTGTGGGCGGAGACATGATTTATCTAGG----- 701  
DB 935 ATGGCTGTGTCTTACAATGCTCCAGTGGGCGGAGGTTGTACTCCCTGGAGGTTCCA 994  
QY 701 CCGTGTACCTCAACGACTCCACTATGACGGACATACCTGGGCGGAGTGCAGTATGGG 760  
DB 995 TCATTTATCTGAATGATTCTGTATGATGGAGCTTGTGGATACAGCATGACAGAGGG- 1054  
QY 761 GTCTGGCGGAGCTGGCAGATGGTGGTGGGCTGGATGACTTTTAGGAGAGTCAAGGAGC 820  
DB 1054 --CTAGGCCAAATGACCGATGGTGTCTGTGGCTGGACGATTTACCCAGACGCCATGAAT 1111  
QY 821 TGGCGGCTGGCCAGCTATGACTATGTGGATGGAGCAACACAGCTTCTCCAGTGGCT 880  
DB 1112 ACCACGTGTGGCGGCGGCTATGACTATGTGGGCTGGGGAACGAGAGTGCACCAATGGCT 1171

QY 881 ATGTGGAGATGGAGTTTGTAGTTTACCGGCTGAGGGCTTCCAGGCTATCAGGTCCACT 940  
DB 1172 ACATTTAGATCATGTTTGAATTTGACCGCATCAGGAATTTTCACTACCATGAAGTCCACT 1231  
QY 941 GTAACAACATGCACACGCTGGAGCGCGTGTGCTGGCGGGGTGAATGTCGTCTCGGC 1000  
DB 1232 GCAACAACATGTTTGTCTTAAAGTGTGAAGATCTTTAAGGAGGTACAGTGTCTCGCT 1291  
QY 1001 GTGGCCCTGCATGGCTGGAGGGGGAGGAGCCATGCGCCACACCTAGGGGGCAACCTGG 1060  
DB 1292 CTG---AAGCCAGTGAAGTGGGAACCTAATGCCATTTCTTCCCTTGTCTGGATGAGC 1348  
QY 1061 GGGACCCAGAGCGCGGCTGTCTCAGTGGCCCTTGGCGGCGTGTGGCTGCTGCTTCTGC 1120  
DB 1349 TCAACCCAGTGTCTGGTGTGTACAGGTGCTCTCCACACCAAGTGGCCAGTGCATCA 1408  
QY 1121 AGTCCGCTCTCTTGGCGGGCGCTTGTACTTCTTCAGGAAATCTCTTCTCATCTCTG 1180  
DB 1409 AGTGTCAATACCATTTTGCAGATACCTGGATGATGTTTCACTGAGATCACCTTCCAATCAG 1468  
QY 1181 ATGTGGTGAACAATTCCTCTCCGCACTGGAGGCACTTCCCGCCAGCCCTGCTGGTGGC 1240  
DB 1469 ATGTGCAATGTACAACAATCTGAAGCCCTGCCCACTCTCTCTA----- 1514  
QY 1241 CGCTGGCGCCACCTCCCAACCACTTACAGACTTGGAGCTGGAGCCGAGAGCCAGCAGC 1300  
DB 1514 -----TGGCACCCACCACTATGATC 1534  
QY 1301 CCGTGGCAACCGGAGGGAGCCCGACCCATCTCTATCTGCTGCTGGTGGCCATCA 1360  
DB 1535 CAATGCTTAAAGTATGATGACAGCAACACTCGGATCTCTGATTGCTGCTGGTGGCCATCA 1594  
QY 1361 TCCTGCTCTCTGCTCATCATTTGCCCTCATGCTCTGGCGGCTGCACTGGCGCAGGCTCC 1420  
DB 1595 TCTTTATCTCTGCGGCATCATTTGTCATCTCTCTGGAGCGAGTCTTGGCAGAAAATGC 1654  
QY 1421 TCAGCAAGGCTGAACGAGGCTGTGGAAGAGGAGTTCACGCTTCACTCTCTGTGCTCTG 1480  
DB 1655 TGGAGAGGCTTCTCGGAGGATGCTGGATGATGAATGACAGTCACTGCTTCCCTGCGCAA 1714  
QY 1481 GGGACATATCTCTATCAACAACCGCCAGGTCTTAGAGAGCCACCCCGCTACAGGAGC 1540  
DB 1715 GTGATCTAGCATGTTCAACAAT-----AACCGCTCT 1747  
QY 1541 CCGGCGCTCTGGGAATCCCGCCCACTCCCTCTCTGTGTCCTCAATGGCTCTGCTGCTGC 1600  
DB 1748 CATCACCTAGTGAACAAGGCTCCAACCTCGACTTACGATCGCATCTTT----- 1795  
QY 1601 TGCTCTCAATCCAGCCTACCGCTCTCTTCTGGCCACTTACGCCCGTCCCGCTCGAGGC 1660  
DB 1795 -----CCCTTCGCTC 1804  
QY 1661 CGGCGCGCCCAACCGCTGGGCCAAACCCACCAACCCAGGCTTACAGTGGGACT 1720  
DB 1805 CTGACTACAGGAGCCATCCAGGCTGATACGAAATCTCCAGAAATTTGCTCCAGGGGAGG 1864  
QY 1721 ATATGGAGCTGAGAACCCAGGCGCGGCTTCTGCCCCACCTCCCGCAGAACAGGCTCC 1780  
DB 1865 AGGAGTCAGGCTGCAGCGGTGTTGTGAAGCCAGTCCAGCCCTGAGCGGCTGAGGGGTGC 1924  
QY 1781 CCCATTATCGGAGCTGACATTTTACCTTGCAGGCGGTCAACCGGGCAACACCTATG 1840  
DB 1925 CCCATATGAGAGGCTGACATAGTGAACCTCCAGAGGTGACAGGAGCAACACATACT 1984  
QY 1841 CTGTGCTGCACTGCCCGCCAGGGGAGTCT---GGGATGGGCGCCCGCCAGAGTGGATTTCC 1897  
DB 1985 CAGTGTCTCGCTTCCAGTGTCTCTCAGGAAAAGATGTGGCTGTGGAGAGTTC 2044  
QY 1988 CTGATCTCCAGTCCCGCTTCAAGAGAGAGTCTGGCGAGGCGCAGTTTGGGAGGCTGCACC 1957  
DB 2045 CCAGGAAACTCTTAACCTTTTCAAGAGAGAGTGGGAGAGGACAGTGTGGGAGGCTTCATC 2104

QY 1958 TGTGTAGGTCGACAGCCCTCAAGATCTGGTACGCTTTGATTTCCCTTAAATGTCGTA 2017  
DB 2105 TCTGTGAAGTGGAGGAATGGAATAATTTTCCCTAGATCTCAGTG 2164  
QY 2018 AGGACACCCCTTCTGGTACGCTGAAGATCTTTACGGCCAGATGCCACCAAGAATGCCA 2077  
DB 2165 CCAACACGCTGCTGCTGGTCTGAAATGCTTCCGAGCAGATGCCAACAAGAATGCCA 2224  
QY 2078 GGAATGATTTCTCTGAAAGAGTGAAGATCATGTCTGAGGCTCAAGGCCCAACATCATTC 2137  
DB 2225 GGAATGATTTCTTAAGAGATAAAGATCATGTCTGCTCAAGGCCCAACATCATTC 2284  
QY 2138 GGCTGTGGGCGTGTGTGACGAGCAGACCCCTCTCCATGATTAATGATCATCATGAGA 2197  
DB 2285 ATCTATTATCTGTGTATCACTGATGACCTCTCTGTATGATCACTGAATATACATGAGA 2344  
QY 2198 ACGGACCTCAACAGCTTCTCAGTCCGACAGCTGGAGGACAGGCGAGGGG 2257  
DB 2345 ATGAGATCTCAATCAGTTCTTTCCGCGCAGAGCCCTTAATCTCTCCAGCG--- 2402  
QY 2258 CCCCTGGGCGGCGGCTGCGCAGGCGGCGCCATCATCAGCTACCCATCTGCTGCATG 2317  
DB 2402 -----ATGACGCACTGTCACTTACACCAATCTGAAGTTTA 2437  
QY 2318 TGGCAGCCAGATCGCTCGGCATGCGCTATCTGGCCACACTCAACTTTGTATCATCGGG 2377  
DB 2438 TGGCTACCAAAATGCTCTGGCATGAAGTACCTTTCTCTTAATTTTGTTCACCGAG 2497  
QY 2378 ACCTGGCCACCGAATGCTCTGCTGAGGAAATTTTACCACCAAAATGATGACTTTG 2437  
DB 2498 ATCTGGCCACAGAACTGTTTATGTTGGAAGAACTACAACTAAGATGACTTTG 2557  
QY 2438 GCATGAGCGGAACTCTATCTGCTGGGACTATTACCTGTGCGAGGCGGCGGCTGCTGC 2497  
DB 2558 GAATGAGCAGAACTGTACAGTGTGACTATTACCGGATCCAGGCGCGGCGAGTCTCC 2617  
QY 2498 CCATCCGCTGATGCTGGGAGTGCATCTCTATGCGGAAGTTTACAGCTGCGAGTGAAG 2557  
DB 2618 CTATCCGCTGATGCTTGTGGAGAGTATCTTGTGGCAAGTTTCACTACACGAAGTATG 2677  
QY 2558 TGTGGCTTTGGTGTGACCTGTGGAGGTGCTGTGATGCTCTGTAGGCGCCAGCCCTTG 2617  
DB 2678 TGTGGCTTTGGGTTGACTTTGTGGGAGTCTTCACTTTTGTCAAGAACAGCCCTATT 2737  
QY 2618 GGCAGCTCACCAGCAGCAGCTCATCGAGAACGCGGGGAGTTCTTCCGGGACAGGCGC 2677  
DB 2738 CCCAGCTGTGATGACAGCTTATTGAGAATACTGAGAGTCTTCCGAGACCAAGGA 2797  
QY 2678 GGCAGGTGTACTGTCCGCGCGCTGCTCCCGCAGGCGCTTATATGAGCTGATGCTTC 2737  
DB 2798 GGCAGACTTACCTCCCTCAACAGCCATTTGTCTGACTCTGTGTATTAAGCTGATGCTCA 2857  
QY 2738 GGTGCTGGAGCGGGAGTCTGAGCAGCGACACACCTTTTCCAGCTGCATCGGTTCT 2795  
DB 2858 GCTGCTGGAGAAGATACGAGAAGCCGCTCCCTCATTTCCAGAAATCCACCTTCTGCT 2915

## RESULT 8

ID Q92523 standard; cDNA to mRNA; 3157 BP.  
AC Q92523;  
DT 26-NOV-1995 (first entry)  
DE Human colonic adenocarcinoma kinase 2 (CK-2) cDNA.  
KW Mammary carcinoma kinase 10; MCK-10; transmembrane receptor; CK-2;  
KW receptor tyrosine kinase; colonic adenocarcinoma kinase 2; ss.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT cds 370..2937  
FT /\*tag= a  
PN W09514089-A.  
PD 26-MAY-1995.  
PF 16-NOV-1994; E03799.  
PR 16-NOV-1993; US-153397.

PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.  
PI Alves.FHE, Ullrich A;  
DR WPI; 95-224055/29.  
DR P-PSDB; R75505.  
PT New nucleic acid encoding CK-2 receptor tyrosine kinase - and  
PT derived vectors, transformed cells, proteins and antibodies, useful  
PT for diagnosis and treatment of proliferative and nervous system  
PT diseases and for screening modulators  
PS Disclosure; Page 73-74; 115pp; English.  
PS A member of the mammary carcinoma kinase 10 (MCK-10) receptor  
CC tyrosine kinase family was identified using a PCR (with two  
CC degenerate oligo primer pools based on conserved sequences of the  
CC kinase domains of receptor tyrosine kinases) and cDNA prep. from  
CC colonic adenocarcinoma RNA. The nt sequence of the novel receptor,  
CC designated CK-2, is given in Q92523 and the deduced AA sequence in  
CC R75505. Analysis of CK-2 nt and AA sequences indicated significant  
CC homology with MCK-10 throughout the extracellular, transmembrane  
CC and intracellular regions. The regions of homology extend into the  
CC N-terminus consensus sequence of the discoidin I like family of  
CC proteins. CK-2 was predominantly found in all stromal cells  
CC whereas MCK-10 expression was strongly confined to neoplastic  
CC cells themselves. Between the two RTKs, the juxtamembrane region  
CC is the region of most extensive sequence divergence. 806 T;  
SQ Sequence 3157 BP; 790 A; 802 C; 759 G;

Query Match 18.3%; Score 666.8; DB 1; Length 3157;  
Best Local Similarity 56.8%; Pred. No. 5e-127;  
Matches 1532; Conservative 0; Mismatches 977; Indels 189; Gaps 9;

QY 110 TGCTGCTGCTTTGGTGGCAAGTGGAGATGCTGACATGAAGGACATTTTGTATCTGCTG 169  
DB 395 TGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 454  
QY 170 AGTGGCGGTATGCTGCTGGCATGAGGACCGACCATCCAGACAGTACATCTGCTT 229  
DB 455 ATGTCGGTATCTCTCTGGCATGTCAGGAGGCGCAGATTCCAGATGAGGACATCAGCT 514  
QY 230 CCAGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 289  
DB 515 CCAGTCAGTGGTTCAGATCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 574  
QY 290 ATGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 346  
DB 575 ATGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 634  
QY 347 TGGATCTCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 406  
DB 635 TTGACTTGCACACCTTCCATTTTATCACTCTGGTGGGACCCAGGCGGCCATGCAGAG 694  
QY 407 GCCTGGCAAGAGTCTCTCCGGAGCTACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 466  
DB 695 GTCATGGCATCGAGTTTGGCCCATGTACAAGATCAATACAGTCGGGATGSCACTCGCT 754  
QY 467 GGATGGCTGGAAGACCGCTGGGGTCAGGAGGTGATCTCAGGCAATGAGGACCTGAGG 526  
DB 755 GGATCTCTTGGCGGAACCGCTCATGGGAAACAGGCTGCTGGATGGAATAGTAACCCCTATG 814  
QY 527 GAGTGTGCTGAAGACCTTGGGCGCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 586  
DB 815 ACATTTTCTTAAGGACTTGGAGCGCCCATTTAGCCAGATTTTCCGGTTTATTCAG 874  
QY 587 GGGCTGACCGGTCTATGAGCGTCTCTCTCGGGTAGAGCTCTATGGCTGCTGCTGAGGG 646  
DB 875 TCACCGACCACTCCATGAATGTGTATGAGAGTGGAGCTTTACGGCTGTGCTGCTGCTG 934  
QY 647 ATGGACTCTCTTACACCGCCCTGTTGGGCGCAGACAAATGTATTTATCTGAGG----- 701  
DB 935 ATGGCTTGGTGTCTTACAATATGCTCCAGCTGGGCGAGCTTTGTACTCCTCGAGGTTCCA 994  
QY 701 CCGTGTACCTCAACGACTCCACCTATGACGACATACCTGCGGCGGCTGCGAGTATGGG 760  
DB 995 TCATTTATCTGAATGATCTGTCTATGATGAGCTGTTGGATACAGCATGACAGAGGG- 1054

QY 761 GTCTGGCCAGCTGGCAGATGGTGTGGTGGCTGGATGACTTTAGGAAGATCAGGAGC 820.  
D 1054 --CTAGGCCAATTGACCGATGTGTCTGGCTGGACGATTTACCCAGACCCATGAAT 1111  
QY 821 TCGGGTCTGGCCAGGCTATGACTATGTGGATGAGAACACACAGCTTCTCCAGTGGCT 880  
D 1112 ACCACGTGTGGCCGGCTATGACTATGTGGCTGGCGAAGAGATGCCACCATGGCT 1171  
QY 881 ATGTGGAGATGGAGTTGAGTTGACCGCTCAGGGCCCTTCAGAGCTATGAGGTCCTACT 940  
D 1172 ACATTGAGATCATGTTGAAATTTGACCGCATCAGGAATTTCACTACCATGAAGGTCCTACT 1231  
QY 941 GTAACAACATGCACAGCTGGAGGCCCGCTCTCCCTGGGGGTGAATGTGCTTTCGGCC 1000  
D 1232 GCAACAACATGTTGCTAAAGGTGAAGATCTTTAAGGAGGTACAGTGTACTTCCGCT 1291  
QY 1001 GTGGCCCTGCCATGGCTGGAGGGGGAGCCCATCGGCACCAACCTAGGGGGCAACCTGG 1060  
D 1292 CTG---AAGCCAGTGAAGTGGGAACCTAATGCCATTTCCCTTCCCTTCTCTGGATGAGC 1348  
QY 1061 GGGACCCAGAGCCGGGCTGTCTCAGTGCCTTGGGGCCGTGTGCTGCTTCTGTC 1120  
D 1349 TCAACCCAGTGTGGTTGTACAGGTGCTCTCCACCCAGATGGCCAGTGGCCATCA 1408  
QY 1121 AGTGGCGCTTCTCTTGGGGGGCCCTGGTACTCTCTCAGCGAAATCTCTTCACTCTCTG 1180  
D 1409 AGTGTCAATACCATTTTGCAGATACCTGGATGATGTTCACTGAGATCACTTCCCAATCAG 1468  
QY 1181 ATGTGTGAACAATTCCTCTCCGGACACTGGGAGGACCTTCCCGCCAGCCCTGGTGGC 1240  
D 1469 ATGCTGCAATGTACAACAACCTCTGAAGCCCTGCCACCTCTCTCTA----- 1514  
QY 1241 CGCCTGGCCCACTCTCCACCAACTTCAGCAGCTTGGAGCTGGAGCCAGAGCCAGCAGC 1300  
D 1514 -----TGACCCCAACACCTATGATC 1534  
QY 1301 CCGTGGCCAAAGCCGAGGGAGCCGACCGCCATCCTCATCGGCTGCCTGGTGGCCATCA 1360  
D 1535 CAATGCTTAAAGTTGATGACAGCAACACTCGGATCCTGATTGGCTGTCTGGTGGCCATCA 1594  
QY 1361 TCTGCTCTCTGCTCATCATGCTGCTTCCCTCATGCTCTGCGGGTGACATGGCGCAGGCTCC 1420  
D 1595 TCTTATCTCTCTGGCCATCATGTTGATCATCTCTGAGGCGCTTCTGGCAAAATGTC 1654  
QY 1421 TCAGCAAGGCTGAAGGGAGGTGTGGAAGAGAGCTGACGGTTTCACTCTCTCTGCTGCTG 1480  
D 1655 TGGAGAGGCTTCTCGGAGATGCTGGATGATGAATGACAGTCAAGCTTTCCTGCGCAA 1714  
QY 1481 GGGACACTATCTCAACAACCCGCGAGTCTTAGAGAGCCACCCCGTACCAAGGAGC 1540  
D 1715 GTGATTTAGCATGTTCAACAAT-----AACCGCTCCT 1747  
QY 1541 CCGGCTCTGGGATCGCCCACTCCGCTCCCTGTGTCCCAATGGCTCTGGGTTGC 1600  
D 1748 CATCACTAGTGAACAAGGGTCCAATCGACTTACGATCGCATCTTT----- 1795  
QY 1601 TGCCTTCCAATCCAGCTACCCGCTCTCTGCGCCACTTACGCCGCTCCCGCTCGAGGCC 1660  
D 1795 -----CCCTTCCGCT 1804  
QY 1661 CGGGCCCCCAACACCGGCTGGGCCAAACCAACCAACCGCCCTACAGTGGGGACT 1720  
D 1805 CTGACTTACCAGGAGCCATCCAGGCTGATACGAAACTCCAGAAATTTGCTCCAGGGAGG 1864  
QY 1721 ATATGAGGCTCAGAGCCAGCGCCCGCTTCTGCCCGCCAGCTCCCGCAGACAGCTTC 1780  
D 1865 AGAGTCAAGGCTGACGCGGTGTGTAAGCCAGTCCAGCCAGTGGCCCTGAGGGGTGC 1924  
QY 1781 CCATTATCCGAGGCTGACATTTGTAACCTGCAGGGGTACACGGGGGCAACACTATG 1840  
D 1925 CCCACTATCGAGGCTGACATAGTGAACCTCCAAGGAGTACAGAGGCAACACATACT 1984

QY 1841 CTGTGCTGCTGCTGCCCCCAGGGGCAGTCT---GGGATGGCCCCCAGAGTGGATTTC 1897  
D 1985 CAGTGCCTGCGCTCACCATGGACTCTCTCAGAAAAAGATGTGGTGTGGAGGATTC 2044  
QY 1988 CTGATCTCGACTCCGCTTCAAGGAGAGCTTGGCAGGGCCAGTTTGGGGAGGTGCACC 1957  
D 2045 CCAGGAACTCTTAACCTTTCAAGAGAAAGCTGGGAGAACAGTGTGGGAGGTTCATC 2104  
QY 1958 TGTGTGAGGTGCACACGCCCTCAAGATCTGGTCACTTGTATTTCCCTTAAATGCGGTA 2017  
D 2105 TCTGTGAAGTGGAGGAATGGAATAATCAAGAACAAAGATTTTCCCTAGATGTCAAGT 2164  
QY 2018 AGGACACCCCTTTCTGCTGAGTGTCAAGATCTTACGGCCAGATGCCAACAAATGCCA 2077  
D 2165 CCAACAGCCTGCTCTGGTGGCTGTGAATGCTCCGAGCAGATGCCAACAAATGCCA 2224  
QY 2078 GGAATGATTTCTTGAAGAGGTGAAGATCATGTCTGAGGTCTAAGGACCCAAACATCTTC 2137  
D 2225 GGAATGATTTCTTGAAGAGATAAAGATCATGTCTCGGCTCAAGGACCCAAACATCTTC 2284  
QY 2138 GGCTGTCTGGCGTGTGTGTCAGGACGACCCCTCTGTCATGATTACTGACTACATGAGA 2197  
D 2285 ATCTATTATCTGTGTATCACTGATGACCCCTCTCTGTATGATCACTGAATACATGAGA 2344  
QY 2198 ACGGCCAGCTCAACACAGTCTCTCAGTGCCTCCACCCAGCTGGAGGACAAAGGAGCGGG 2257  
D 2345 ATGGAGATCTCAATCAGTTCTTTCCGCCACGAGCCCTTAATCTTCTCCAGCG--- 2402  
QY 2258 CCCTGGGAGCGGCGAGGTGCGCAGGGCCCAACCATACAGTACCCAAATGCTGCTCATG 2317  
D 2402 -----ATGTACGCACTGTCACTGTACCAATCTGAAGTTTA 2437  
QY 2318 TGGCAGCCAGATCGCTCGGCATGGCTATCTGTGGCCACACTCACTTTGTACATCGGG 2377  
D 2438 TGGTACCCAAATTTGCTCTGGCATGAAGTACCTTTCTCTCTTAATTTGTTCACCGAG 2497  
QY 2378 ACCTGGCCAGCGGAATCTGCTAGTTGGGAAAAATTTACCATCAAAATTCGACAGCTTG 2437  
D 2498 ATCTGGCCACACGAACTGTTTAGTGGTAAAGAACTACAAATCAAGATAGCTGACTTG 2557  
QY 2438 GCATGAGCCGGAACCTCTATGCTGGGAGTATTTACGCTGTGAGGCGCGGAGTCTGCTC 2497  
D 2558 GAATGAGCAGGAACCTGTACAGTGGTGAATTTACCGGATCCAGGCGCGGAGTCTCTC 2617  
QY 2498 CCATCGCTGAGTGGCTGGGAGTGCATCTCATGGGAGTTCACAGTGCAGAGTGCAGC 2557  
D 2618 CTATCCCTGGATGCTCTGGAGAGTATCTGCTGGCAAGTTCACAGCAAGTGTG 2677  
QY 2558 TGTGGCCCTTTGGTGTGACCCCTGTGGAGGTGCTGATGCTGTAGGGCCAGCCCTTTG 2617  
D 2678 TGTGGCCCTTTGGGTTTACTTTGTGGAGACTTTTCACTTTGTCAAGAACAGCCCTATT 2737  
QY 2618 GGAGCTCACGAGCAGAGGTGATCGAGAACCGGGGGAGTCTTCCGGGACGAGGCC 2677  
D 2738 CCAGCTGTGATGATGAGAGTATTGAGATACTGAGAGTCTTCCGAGACCAAGGGA 2797  
QY 2678 GGAGGTGTACCTGCTCCGCGCCCTGCTGCGCGAGGCGCTATATGAGTGTGATCTTC 2737  
D 2798 GGAGACTTACCTCCCTCAACAGCCATTTGCTGACTCTGTGTATGATGATGATCTCA 2857  
QY 2738 GGTGCTGGAGCGGAGTGTGAGCAGCGACCCCTTTTCCAGCTGCTATCGGTTCT 2795  
D 2858 GCTGCTGGAGAGAGATACGAGAACCGTCCCTCATTTCCCAAGAAATCCACTTCTGCT 2915

## RESULT 9

T93784

ID T93784 standard; cDNA; 3157 BP.

AC T93784;

DT 16-FEB-1998 (first entry)

DE CCK-2, a human mammary carcinoma kinase 10 MCK-10) family member, cDNA.

KW Mammary carcinoma kinase; MCK-10; CCK-2; receptor tyrosine kinase;

proliferative disease; cancer; MCK-10 activity; aberrant expression; ds.

OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT CDS 370..2934  
 FT /\*tag= a  
 PN US5677144-A.  
 PD 14-OCT-1997.  
 PF 08-NOV-1994; 336343.  
 PR 16-NOV-1993; US-153397.  
 PA (ALVE/) ALVES F H E.  
 PA (ULLR/) ULLRICH A.  
 PI Alves FHE, Ullrich A;  
 DR WPI: 97-511869/47.  
 DR P-PSDB; W34671.  
 PT Truncated receptor tyrosine kinase CKK-2 - and nucleic acid coding  
 PR for it, useful for cancer diagnosis  
 PS Claim 9; Fig 3; 70pp; English.  
 CC The present sequence represents the cDNA sequence of human CKK-2, a  
 CC member of the mammary carcinoma kinase 10 (MCK-10, T93785) family of  
 CC receptor tyrosine kinases. Expression of CKK-2 is associated with  
 CC proliferative diseases such as cancer. The CKK-2 gene was identified by  
 CC PCR and a cDNA prepared from colonic adenocarcinoma RNA. CKK-2 is  
 CC expressed in a wide variety of cancer cell lines and tumour tissue. The  
 CC CKK-2 nucleic acids can be used for diagnostic purposes to detect  
 CC aberrant expression of CKK-2 genes. Engineered cell lines, containing  
 CC recombinant vectors with the present sequence, are useful for producing  
 CC infectious retroviral particles. The cell lines may also be used to  
 CC evaluate and screen drugs involved in CKK-2 activation and regulation.  
 SQ Sequence 3157 BP; 790 A; 802 C; 759 G; 806 T;

Query Match 18.3%; Score 666.8; DB 1; Length 3157;  
 Best Local Similarity 56.8%; Pred. No. 5e-127;  
 Matches 1532; Conservative 0; Mismatches 977; Indels 189; Gaps 9;

QY 110 TGCTGCTGCTCTGGTGGCAAGTGGAGATGCTGACATGAGGAGACATTTTATCTCTGCGCA 169  
 DB 395 TGGTGTGTTCTCTGCTGCTGCTATCTTTAGTGTCTGCAAAAGCTCAGGTTAATCCAGCTA 454  
 QY 170 AGTCCGCTATGCTGCTGGGATGAGGAGCGGACCATCCAGAGAGTGGATCTCTGCTT 229  
 DB 455 TATCCGCTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 514  
 QY 230 CCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 289  
 DB 515 CCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 574  
 QY 290 ATGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 346  
 DB 575 ATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 634  
 QY 347 TGGATCTACACGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 406  
 DB 635 TTGACTTGACACCTCCATTTTATCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 694  
 QY 407 CCTTGGGAAGAGTCTCTCCGAGCTACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 466  
 DB 695 GTATGGCATCGATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 754  
 QY 467 GGATGGCTGGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 526  
 DB 755 GGATCTCTGGCGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 814  
 QY 527 GAGTGGTCTGAAGAGCTTGGGCCCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 586  
 DB 815 ACATTTTCTTAAGAGCTTGGAGCGCCCATTTGAGCCAGATTGTCGGGTTTCAATCCAG 874  
 QY 587 GGGCTGACCGGCTCATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 646  
 DB 875 TCACCGACCACTCCATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 934  
 QY 647 ATGACCTCTGCTTACACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 701  
 DB 935 ATGCTTGGTCTTACAACTCTCCAGCTGGGAGCAGTTTGTACTCCCTTGGAGGTTCCA 994

QY 701 CCCTGTACCTCAACGACTTCCACCTATGACGGACATACCGTGGCGGACTGACAGTATGGG 760  
 DB 995 TCATTTATCTGAATGATTCTGTCTATGATGAGAGCTGTTGGATACAGCATGACAGAGGG - 1054  
 QY 761 GTCGTGGCCAGCTGGCAGATGGTGTGGTGGGCTGGATGACTTTAGGAAGAGTACAGGAGC 820  
 DB 1054 --CTAGGCCAATTTACCGATGGTGTGTGCTGGCCTGGAGATTTTCCACCCAGACCATGAAT 1111  
 QY 821 TCGCGGCTCGCCAGGCTATGACTATGTTGGATGAGCAACACACAGCTTCTCCAGTGGCT 880  
 DB 1112 ACCAGCTGTGGCCGCTATGACTATGTTGGCTGGCGAAGAGAGTGGCCACCAATGGCT 1171  
 QY 881 ATGTGGAGATGAGTGTGAGTTTACCGGCTGAGGGCTTCTCCAGGCTATGACAGTCCACT 940  
 DB 1172 ACATTTGAGATCATGTTTGAATTTGACCGCATCAGAAATTTTCACTTACCATGAAGTCCACT 1231  
 QY 941 GTAACAACATGACACGCTGGGAGCCGCTGTGCTGGCGGGTGGATGTGCTTCCGGC 1000  
 DB 1232 GCACAACATGCTTTGCTAAAGGTGTGAAGATCTTTAAAGAGGTACAGTGTCTACTTCCGCT 1291  
 QY 1001 GTGGCCCTGCCATGCTGCTGGGAGGAGCCATGCGCCACCACTAGGGGGCAACCTGG 1060  
 DB 1292 CTG---AAGCCAGTGTGGGAACCTATGCAATTTCTTCCCTTGTCTCGATGAGC 1348  
 QY 1061 GGGACCCAGAGCCCGGCTGTCTCAGTGGCCCTTGGCGGCGTGTGGCTGCTTCTG 1120  
 DB 1349 TCAACCCAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1408  
 QY 1121 AGTGGCGTCTCTTTCGGGGGCTGCTTCTTTCAGGGAATCTCTTCTTCTCTCTG 1180  
 DB 1409 AGTGTCAATATACATTTTTCAGATACCTTGGATGATGTTTTCAGTGTGATCCTTCAATCAG 1468  
 QY 1181 ATGTGGTCAACAATTTCTCTCCGCTGCTGGAGGACCTTCCCGCCAGCCCTTGGTGGC 1240  
 DB 1469 ATGTGCAATGTACAACTCTGAGCCCTGCGCCACTCTCTTA----- 1514  
 QY 1241 CGCTTGGCCACTTCCCACTTACAGAGCTTGGAGCTGGAGCCAGAGCCAGCAGC 1300  
 DB 1514 -----TGGCCACCAACCTATGATC 1534  
 QY 1301 CCGTGGCCAAAGCCGAGGAGCCGACCGCATCTCTATCGGCTGCTGCTGGCCATCA 1360  
 DB 1535 CAATGCTTAAAGTTGATGACAGCAACACTCGGATCTGCTGCTGCTGCTGCTGCTGCTGCT 1594  
 QY 1361 TCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1420  
 DB 1595 TCTTTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1654  
 QY 1421 TCAGCAAGCTGAAGGAGGCTTGGAGAGAGCTGACGGTTTCACTCTCTCTGCTGCTGCT 1480  
 DB 1655 TGGAGAAGGCTTCTCGGAGGATGCTGGATGATGAATGACAGTCACTGCTGCTGCTGCTGCT 1714  
 QY 1481 GGGACACTATCTCATCAACAACCGCCAGCTCTAGAGAGCCACCCCTGACAGGAGC 1540  
 DB 1715 GTGATTTAGCATGTTCAACAAT-----AACCGCTCT 1747  
 QY 1541 CCGGCTCTGCTGGGAATCCGCGCCACTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1600  
 DB 1748 CATCACTTGTGAACAAGGTTCCAACTGCTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1795  
 QY 1601 TGCTCTCAATCCAGCTTACCGCTCTCTTCTGGGCACTTACGCGCTGCTGCTGCTGCTGCT 1660  
 DB 1795 -----CCCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1804  
 QY 1661 CGGGCCCCCAGACCCGCTGGGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1720  
 DB 1805 CTGACTACAGAGGCTTCCAGGCTGATACGAACTCCCAAGATTTGCTCCAGAGGAGG 1864  
 QY 1721 ATATGAGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1780  
 DB 1865 AGGAGTCAAGGCTGAGCGGTGTTGTGAAGCCAGTCCAGCCAGTGGCCCTGAGGGGTGC 1924

QY 1781 CCCATTATCGGAGGCTGACATTTTACCTGTCAGAGGCGTCACCGGGGCAACACCTATG 1840  
 Db 1925 CCCACTATGACGGCTGACATAGTAACTTCAAGGAGTACAGAGGAGGCAACACATACT 1884  
 QY 1841 CTGTGCTGACACTCCGCCAGGSCAGTC--GGGATGGCCCCCAGAGTGGATTTC 1897  
 Db 1985 CAGTGCCTGGCTACCATGACCTGCTCAGGAAAGATGTGGCTGTGGAGGATTC 2044  
 QY 1998 CTCGATCTGACCTCCGCTTCAAGGAGAAGCTTGGCGAGGCGCAGTTTGGGGAGGTGCAC 1957  
 Db 2045 CCAGGAAACTCCTAACTTCAAGAGAAGCTGGGAGAAGCAGATTGGGGAGGTTCATC 2104  
 QY 1958 TGTGTAGGTGCGACGCGCTCAAGATCTGTGAGTCTTGTATTTCCCTTAATGTGCGTA 2017  
 Db 2105 TCTGTGAAGTGGAGGAATGGAAATTTCAAGACAAAGATTTTGGCCTAGATGTCA 2164  
 QY 2018 AGGACACCTTCTGCTGTAGCTGTCAAGATCTTACGGCCAGATGCCACCAAGATGCCA 2077  
 Db 2165 CCAACGAGCTGCTGCTGTGCTGTGAATGCTCCAGCAGATGCCAAGAAATGCCA 2224  
 QY 2078 GGAATGATTTCTGAAAGAGGTGAAGATCATGTCGAGGCTCAAGGACCCCAACATCATTC 2137  
 Db 2225 GGAATGATTTCTTAAAGAGATGAAGATCATGTCGCTCAAGGACCCCAACATCATTC 2284  
 QY 2138 GGCTGTGGGGGTGTGTGAGACACCCCTCTGTCATGATTAATGACTACATGAGA 2197  
 Db 2285 ATCTATTATCTGTGTATCATCTGATGACCTCTCTGTATGATCACTGAATACATGAGA 2344  
 QY 2198 ACGCGACCTCAACAGCTTCTCAGTCCCGCCAGCTGGAGGACAGGAGCGGAGGG 2257  
 Db 2345 ATGAGATCTCATCAGTTTCTTCCCGCCAGGCCCCCTTAATCTTCTCCAGCG-- 2402  
 QY 2258 CCCTGGGGACGGCAGGCTGCGAGGGGCCACCATCATGCTCAATGCTCTGCTCATG 2317  
 Db 2402 -----ATGACGCACTGTGCTTACACCAATCTGAAGTTTA 2437  
 QY 2318 TGGCAGCCAGATCGCTCCGGCATCGGCTATCTGGCCACACTCACTTTGTACATCGGG 2377  
 Db 2438 TGGTATCCCAAAATGGCTCTGGATGAAGTACCTTTCTCTTAATTTTGTTCACCGAG 2497  
 QY 2378 ACCTGGCCACCGGAAGTCTAGTTGGGAAATTTACCAATCAAAATCGGAGCTTGG 2437  
 Db 2498 ATCTGGCCACAGAACTGTTAGTGGGTAAAGAACTACAAATCAAGATGACTGTG 2557  
 QY 2438 GCATAGCCGGAACCTCTATGCTGGGACTATTACCGTGTGCGAGGCGCGGAGTGTCT 2497  
 Db 2558 GAATGACGGAACCTGTACAGTGGTACTATTACCGGATCCAGGCGCGGAGTGTCT 2617  
 QY 2498 CCATCCGCTGGATGGCTGGAGTGCATCTCATGGGAAGTTACAGCTGGAGTGCAG 2557  
 Db 2618 CTATCCGCTGGATGCTTGGGAGATATCTTGTGGGCAAGTTCACTACAGCAAGTGTG 2677  
 QY 2558 TGTGGGCTTTGGTGTGACCTGTGGAGGTGCTGATGCTGTAGGGCCAGCCCTTTG 2617  
 Db 2678 TGTGGGCTTTGGGTACTTTTGGGAGCTTTACCTTTTGAAGAACGCCCTATT 2737  
 QY 2618 GGCAGCTCACCGACGAGCTCATCGAAGACCGGGGAGTTCTTCCGGGACCGAGGCC 2677  
 Db 2738 CCCAGCTGTGAGATGAACAGTTATTGAGAAATCTGGAGATTTCTCCGAGACCAAGGA 2797  
 QY 2678 GGCAGGTGTACCTGTCCCGCCGCTGCTGCGCCGAGGCGCTATATGAGCTGTAGCTTC 2737  
 Db 2798 GGCAGACTTACCTCCCTCAACAGGCAATTTGTCTGACTCTGTGTATAGCTGATGCTCA 2857  
 QY 2738 GGTGCTGGAGCGGGAGTCTGAGCAGGACGACCCCTTTTCCAGCTGCATCGTTCTCT 2795  
 Db 2858 GCTGCTGGAGAAGATACCAAGAACCGTCCCTCATTTCCAGAAATCCACCTTCTCT 2915

RESULT 10  
 V48292  
 ID V48292 standard; cdna; 3096 BP.  
 AC V48292;

16-NOV-1998 (first entry)  
 Discoidin domain receptor 2 gene.  
 KW Discoidin domain receptor; transformation; metastasis; collagen; ss;  
 KW Cleidocranial dysplasia; Sicker syndrome; extracellular matrix; MMP-1.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT CDS 354..2921  
 FT /\*tag= a  
 FT /product= "Discoidin domain receptor"  
 FT 354..416  
 FT /\*tag= b  
 FT mat\_peptide 417..2918  
 FT /\*tag= c  
 WO9834954-A2.  
 PD 13-AUG-1998.  
 PF 05-FEB-1998; CA0093.  
 PR 06-FEB-1997; US-041578.  
 PA (MOUN) MOUNT SINAI HOSPITAL CORP.  
 PI Pawson A, Vogel W;  
 DR WPI; 98-447168/38.  
 DR P-PSDB; W7114.  
 DR Novel ligands of discoidin domain receptor tyrosine kinase,  
 especially collagen - useful for treating e.g. metastasis,  
 PT cleidocranial dysplasia or Sicker syndrome  
 PS Disclosure; Fig 22a; 115pp; English.  
 CC The DDR can be used to identify and evaluate substances which affect DDR  
 CC receptor tyrosine kinase signalling pathways in the cell. Compounds  
 CC which modulate such signalling pathways can be used to alter  
 CC structural or functional deregulation of collagens, e.g. Cleidocranial  
 CC dysplasia or Sicker syndrome, conditions requiring modulation of  
 CC extracellular matrix synthesis, degradation or remodelling, or to treat  
 CC conditions needing modulation of MMP-1 expression such as wound healing.  
 SQ Sequence 3096 BP; 762 A; 791 C; 752 G; 791 T;

Query Match 18.3%; Score 666.8; DB 1; Length 3096;  
 Best Local Similarity 56.8%; Pred. No. 5e-127;  
 Matches 1532; Conservative 0; Mismatches 977; Indels 189; Gaps 9;

QY 110 TGCTGCTCTTGGTGGCAAGTGGAGATGCTGACATGAAGGGACATTTGATCCTGCCA 169  
 Db 379 TGGTGTCTTCTCTGCTGCTATCTTGTAGTTCTTCAAAAAGCTCAGGTTAATCCAGCTA 438  
 QY 170 AGTGGCGCTATGCTGGCATGCAGGACCGACCATCCAGACAGTGACATCTCTCTT 229  
 Db 439 TATGCGGTATCTCTGGGCAATGTCAGAGCCAGATTCAGATGAGGACATCACAGCTT 498  
 QY 230 CCAGTCTGCTGATTCCTACTGCCCGCCACAGCAGGTTGGAGAGCAGTGACGCGGG 289  
 Db 499 CCAGTCACTGGTCCAGAGTCCACAGCTGCCAAATATGGAAGCTGGACTCAGAAAGGGG 558  
 QY 290 ATGGGGCTGTGCTGGCGGAGGTCGGTGTTCCTCCAGGA---GGAGAGTACTTGCAGG 346  
 Db 559 ATGGAGCCTGTGCTGGCTGAGATTCCAGTGGAAACCTGTGATGACCTGAAGAGTTTCTGCAGA 618  
 QY 347 TGGATCTACAACGACTGCACCTGTGCTGTGGCCACCGAGGACGCGCATCCCGGG 406  
 Db 619 TTGACTTGCACACCTCCATTTTATCACTCTGTGGGAGCCAGGGGCGCATCCAGAG 678  
 QY 407 GCCTGGGCAAGAGTCTTCCGGAGCTACCGGCTGCGTTACTCCCGGATGGTCCGCGCT 466  
 Db 679 GTCATGCGATCGAGTTTGGCCCCCATGTACAAGATCAATTACAGTCGGAATGGCACTCGCT 738  
 QY 467 GGATGGGCTGGAAGGACCGCTGGGTCAGGAGTGTATCAGGCAATGAGACCTCAGG 526  
 Db 739 GGATCTCTTGGCGAAGCCGTCATGGGAAACAGGTGCTGGATGGAATAGTAACCCCTATG 798  
 QY 527 GAGTGGTCTGAAGGACCTTGGGCCCCCATGTTGTCGCGACTGGTTCGCTTCTACCCCC 586  
 Db 799 ACATTTTCTTAAGAGACTTGGAGCCGCCCATTTAGCCAGATTTGTCCGGTCTATCCAG 858  
 QY 587 GGGCTGACCGGGTCTATGAGCGTCTGTCTGGGGTAGAGCTCTATGCTGCTCTGGAGGG 646



Db 859 TCACCGACACCTCCATGAATGTTATGAGAGTGGAGCTTTACGGCTGTGCTGGCTAG 918  
QY 647 ATGGACTCTCTTACACCGCCCTGTGGGACACAAATGTAATTTATCTGAGG-----701  
Db 919 ATGGCTTGTGTTTACATGCTCCAGCTGGGACAGCTTTGTACTCCCTGGAGTTCCA 978  
QY 701 CCGTGTACCTAACGACTCCACTATGACGGACATACCGTGGGGAGTCCAGTATGGG 760  
Db 979 TCATTTATCTGAATGTTCTTATGATGGAGCTGTGGATACAGCATACAGAAAGG-1038  
QY 761 GTCTGGGACAGCTGCAGATGTTGTTGGGCTGGATGACTTTAGGAAGAGTCAGAGC 820  
Db 1038 --CTAGGCCAATGTACCGATGTTGCTGGCTGGAGATTTACCCAGACCCATGAAT 1095  
QY 821 TCGGGCTGTGGCAGGCTATGACTATGTTGGATGGAGCAACACAGCTTCTCCAGTGGCT 880  
Db 1096 ACCAGTGTGGCCGCTATGACTATGTTGGCTGGCGAAGAGAGTCCACCAATGGCT 1155  
QY 881 ATGTGGAGATGAGTTTGAAGTTGACCGGCTGAGGGCTTCCAGGCTATGAGGTCCACT 940  
Db 1156 ACATTGAGATCATGTTGAATTTGACCGCATCAGGAATTTCACTACCATGAAGTCCACT 1215  
QY 941 GTAAACATGCACAGCTGGAGCCGCTCTGCTGGGGGTGGAATGTCGCTCCGGC 1000  
Db 1216 GCAACAACATGTTGCTAAGGTGTGAAGATCTTTAAGGAGGTACAGTGTACTTCCGCT 1275  
QY 1001 GTGGCCCTGCCCTGGAGGGAGCCCATCGGCACAACTAGGGGGCAACCTGG 1060  
Db 1276 CTG---AAGCCAGTGAAGTGAAGTATGCTTCCCTTCTCCCTGGATGAG 1332  
QY 1061 GGAACCCAGAGCCGGGCTGTCTCAGTGGCCCTTGGGGGGGCTGTGCTCGCTTTGCG 1120  
Db 1333 TCAACCCAGTGTGGTGTGTCAGGGTGTCTCCACCGAGTGGCCAGTGCATCA 1392  
QY 1121 AGTGGCGCTTCTCTTGGGGGCGCTGTTACTTCTCAGCGAATCTCTCATCTCTG 1180  
Db 1393 AGTGTCAATACCATTTTGCAGATACCTTGGATGATGTTGAGTGAGATCACTTCCAAATCAG 1452  
QY 1181 ATGTGTGAACAATCTCTCGGCACTGGAGGCACTTCCCGGACCCGCTGGTGGC 1240  
Db 1453 ATGCTCAATGTACAACTCTGAAGCCCTGCCACCTCTCTA-----1498  
QY 1241 CGCCTGGCCACCTCCCAACCACTTCAGCAGCTTGGAGCTGGAGCCAGAGCCAGCAGC 1300  
Db 1498 -----TGCAACCCACAACCTATGATC 1518  
QY 1301 CCGTGGCCAAAGCCGAGGGAGCCGACCGCATCTCATCGCTGCGCTGGTGGCCATCA 1360  
Db 1519 CAATGCTTAAAGTTGATGACAGCAACACTCGGATCCTGATGGCTGCTTGTGGCCATCA 1578  
QY 1361 TCCTGCTCCTGCTCATCATTTGGCCCTCATGCTTGGCGGCTGCACCTGGGCGAGGCTCC 1420  
Db 1579 TCTTTATCTCTGGCCATCATTTGATCATCTCTCTGGAGGCACTTGGCAAAATGC 1638  
QY 1421 TCAGAAGGCTGAACGGAGGTGTTGGAAGAGAGCTGACGTTTCACTCTCTGTCCTG 1480  
Db 1639 TGGAGAAGGCTTCTCGGAGGATGCTGGATGATGAATGACATGACATGAGCTTCCCTGCCAA 1698  
QY 1481 GGGACACTCTCTCAACAACCCGAGGTCCTTAGAGAGCCACCCCGTACAGGAGC 1540  
Db 1699 GTATTCTAGCATGTTCAACAAT-----AAGCCGCTCT 1731  
QY 1541 CCGGGCTGTGGGAATCGCCCACTCGCTGCTCCCTGTGTCGCCAATGGCTCTGGGTTGC 1600  
Db 1732 CATCACTAGTGAACAGGTCCTCACTCGACTTACGATCGCATCTT-----1779  
QY 1601 TGTCTCAATCAGGCTACGGCTCTCTTGGCCACTTTACCCCGTCCCGCTCGAGGCC 1660  
Db 1779 -----CCCTTCGCC 1788  
QY 1661 CGGGCCCCCACACCGCTGGGCAAAACCCACCAACCCAGGCTACAGTGGGACT 1720

Db 1789 CTGACTACAGGAGCCATCCAGGCTGATACGAAAACTCCAGAAATTTGCTCCAGGGAGG 1848  
QY 1721 ATATGAGCCTGAGAGCCAGCGCCCCGCTTCTGCCCCACCTCCCCAGAACAGGTCC 1780  
Db 1849 AGGAGTCAGGCTCAGCGGTGTTGTAAGCCAGTCCAGCCAGTGGCCCTGAGGGGTGC 1908  
QY 1781 CCAATTTATGCCAGGCTGACATTTGTTACCTCGAGGGCTCACCAGGGGCAACACCTATG 1840  
Db 1909 CCACATATGACAGGCTGACATAGTGAACCTCAAGAGTGCAGAGGCAACACATCT 1968  
QY 1841 CTGTGCTGCACATGCCCCAGGGGAGTC---GGGATGGGCCCCCAGAGTGGATTTCC 1897  
Db 1969 CAGTGGCTGCCGTACCATGGACCTGCTCTCAGAAAAGATGTGGCTGTGGAGGATTC 2028  
QY 1898 CTCGATCTCGACTCCGCTTCAAGGAGAAGCTTGGCGAGGGCCAGTTTGGGAGGTGCAC 1957  
Db 2029 CCAGGAAACTCCTTAACCTTCAAGAGAAGCTGGGAGAAGACAGTTTGGGAGGTTCATC 2088  
QY 1958 TGTGTGAGTTCGACACCCCTCAAGATCTGTCAGTCTTCAATTTCCCTTAAATGTCGTA 2017  
Db 2089 TCTGTGAAGTGGAGGAATGGAATAATTCAAAGACAAGATTTGCCCCAGATGTCAGTG 2148  
QY 2018 AGGACACCCCTTGTCTGTAGTGTCAAGATCTTACGGCAGATGACCAACAAGATGCCA 2077  
Db 2149 CCAACAGCTGTCTGTGGCTGTGAAATGCTCCGACAGATGCCAACAAAGATGCCA 2208  
QY 2078 GGAATGATTTCTTAAAGAGTGAAGATCATGTGAGGCTCAAGGACCCAAACATCATTC 2137  
Db 2209 GGAATGATTTCTTAAAGAGATAAGATCATGTCTGGCTCAAGGACCCAAACATCATCC 2268  
QY 2138 GGCTGTGGCGTGTGTGTCAGGACGACCCCTCTGCAATGATTACTGACTACATGGAGA 2197  
Db 2269 ATCTATTATCTGTGTATCACTGATGACCCCTCTCTGTATGATCACTGAATACATGGAGA 2328  
QY 2198 ACCGCACTTCAACCACTTCTCAGTGCACCCAGCTGGAGGACAAGGAGGAGGGG 2257  
Db 2329 ATGGAGATCTCAATCAGTTTCTTCCCGCCACAGAGGCCCTTAATTTCTCTCCAGCG---2386  
QY 2258 CCCCTGGGAGCGGCGAGGCTGCGCAGGGCCCACTACAGTACCCAAATGCTGCTGCATG 2317  
Db 2386 -----ATGTAGGCACTGTCACTGATACCAATCTCAAGTTTA 2421  
QY 2318 TGGCAGCCAGATCGCTCCGGCATCGGCTATCTGCGCACACTCACTTTGTACATCGGG 2377  
Db 2422 TGGCTACCCAAATTTGGCTGTGGCATGAAGTACCTTTCTCTCTTAATTTGTTCACGAG 2481  
QY 2378 ACCTGGCCACCGCACTGCTAGTTGGGAAAATTTCAACCATCAAAATTCGCGAGCTTG 2437  
Db 2482 ATCTGGCCACACCAACTGTTTGTGGTAAAGAACTACAACTACAGATAGTGTGCTTG 2541  
QY 2438 GCATGAGCCGAACCTCTATGCTGGGAGTATPACGCTGTGAGGCGCGGAGTCTGCTG 2497  
Db 2542 GAATGAGCAGAACTGTACAGTGTGACTATTACCGGATCCAGGGCGGCGAGTCTCC 2601  
QY 2498 CCATCCGCTGGATGGCTGGAGTGCATCTCATGGGAGTTCAGACTGGAGGTGAGC 2557  
Db 2602 CTATCCGCTGGATGCTTGGGAGATCTTGTCTGGGAAGTTCACTACAGAAAGTGTG 2661  
QY 2558 TGTGGGCTTTGTGTGACCTGTGGAGGTGCTGATGCTGTAGGGCCAGCCCTTTG 2617  
Db 2662 TGTGGGCTTTGGGTTACTTTGTGGAGACTTTTACCCTTTGTCAAGAACAGCCCTATT 2721  
QY 2618 GGCAGCTTCCAGCAGCAGGCTCATCGAAGACCGGGGAGTTCTTCCGGGACAGGGCC 2677  
Db 2722 CCCAGCTGTGAGATGAACAGGTTATTGAGAATACTGAGAGTCTTCCGAGCAAGGGA 2781  
QY 2678 GGCAGGTGTACCTGTCCCGCCGCTGCTGCGCCGAGGCGCTATATGAGTGTATCTTC 2737  
Db 2782 GGCAGACTTACCTCCCTCAACAGCCATTTGCTGACTCTGTGTATAAGCTGATGCTCA 2841  
QY 2738 GGTGCTGGAGCCGAGTCTGAGCAGCAGCACCCTTTTCCAGCTGCATCGGTTCT 2795  
Db 2842 GCTGCTGGAAGAGATACCAAGAACCGTCCCTCAITCCCAAGAAATCCACCTTCTGCT 2899



## RESULT 11

Q64158  
ID Q64158 standard; DNA: 2128 BP.  
AC Q64158;  
DT Q3-FEB-1995 (first entry)  
DE Partial coding sequence of tyrosine kinase receptor protein.  
KW Tyrosine kinase; receptor; proto-oncogene; trk; detection;  
KW diagnosis; antibody; treatment; tumour; antisense; ss.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT 1..1953  
FT cds  
FT /tag= a  
FT /product= Partial sequence of protein tyrosine  
FT kinase.  
FN DE4239817-A.  
PD 01-JUN-1994.  
PF 26-NOV-1992; 239817.  
PR 26-NOV-1992; DE-239817.  
PA (CHEM-) CHEMOTHERAPEUTISCHES FORSCHUNG.  
PI Holtrich U, Ruebsamen-waigmann H, Strebhardt K;  
DR WPI; 94-184380/23.  
DR P-PSDB; R54089.  
PT New protein tyrosine kinase and related nucleic acid - vectors,  
PT transformed cells, etc., useful for diagnosis and treatment of  
PT tumours  
PS Claim 4; Page 8; 9pp; German.  
CC The gene is related to the trk proto-oncogene. Antibodies against  
CC the encoded polypeptide are useful for diagnosis and for the  
CC treatment of tumours. The antibodies may also be radiolabelled or  
CC coupled to a cytotoxin for destruction of cancer cells. Antisense  
CC nucleic acid can be used to inhibit gene expression.  
SQ Sequence 2128 BP; 532 A; 556 C; 504 G; 536 T;

Query Match 13.2%; Score 481.4; DB 1; Length 2128;  
Best Local Similarity 56.1%; Pred. No. 2.5e-89;  
Matches 1176; Conservative 0; Mismatches 741; Indels 180; Gaps 7;

QY 702 CGTGTACCTACAGACTCCACTATGACGGAGATACCGTGGCGGACTGCGAGTATGGGG 761  
Db 12 CATTATCTGATGATTCTGTATGATGGAGCTGTGGATACAGCATGACAGAGGG-- 70  
QY 762 TCTGGCGAGCTGGCAGATGTTGGTGGGCTGGATGATGATTTAGGAAGAGTCAGGAGCT 821  
Db 70 -CTAGGCCAATTGACCGATGGTGTCTGTGGCTGGACGATTTACCCAGAGCCCATGAATA 128  
QY 822 GCGGCTCTGGCCAGCTATGACTATGTGGGATGGAGCAACACAGCTTCTCCAGTGGCTA 881  
Db 129 CCACGTGTGGCCGCTATGACTATGTGGGCTGGCGAAGAGAGTGCACCAATGGCTA 188  
QY 882 TGTGAGATGAGATTGAGTTGACCGGCTCAGGCGCTTCAGGCTATGACAGGTCCACTG 941  
Db 189 CATTGAGATCATGTTGATTTGACCGCATCAGGATTTCACTACCATGAGGTCCACTG 248  
QY 942 TAAACATATGACAGCTGGGAGCCGCTCTGCTGGGCGGTGGAATGTCCCTTCGCGG 1001  
Db 249 CAACAACATGTTTGAAGGTGTGAAGATCTTTAAGGAGGTACAGTGTCTACTTCCGCTC 308  
QY 1002 TGGCCCTCCATGGCTGGAGGGGAGCCCATGCGCCACACACCTAGGGGCAACCTGGG 1061  
Db 309 TG---AAGCAGTGTAGTGGGAACCTAATGCCATTTCTTCCCTTGTCTGGATGAGT 365  
QY 1062 GGACCCAGAGCCCGGCTGTCTAGTCCCTTGGCGGCGGTGTGGCTCGCTTCTGCA 1121  
Db 366 CAACCCAGTGTCTGGTTGTACAGGTGCTTCCACCCAGCATGAGGAGTGGCCATCAA 425  
QY 1122 GTGCGCTTCTCTTTGGGGGCGCTGTCTACTCTTACAGGAATCTCTTCATCTCTGA 1181  
Db 426 GTGTCAATACCAATTTTGCAGATACCTGTGATGTTTCACTGAGATCATCTTCCCAATCAGA 485  
QY 1182 TGTGTGACAAATTCCTCTCCGGCACTGGGAGGACCTTCCCGGAGCCCTCTGGTGGCC 1241

Db 486 TGCTGAATGTACAACAACCTCTGAAGCCCTGCCACCTCTCC----- 528  
QY 1242 GCCTGCCCCACCTCCACCACCACTTCAGCAGCTTGGAGCTGGAGCCAGAGGCCAGCAGCC 1301  
Db 528 -----TATGGCACCACCAACCTATGATCC 551  
QY 1302 CGTGGCAAGCCCGAGGGAGGCCGACCCCATCTCTATCGGCTGCTGGTGGCCATCAT 1361  
Db 552 AATGCTTAAAGTTGATGACAGCAACTCGGATCTGATGGCTGCTGGTGGCCATCAT 611  
QY 1362 CTTGCTCTCTGCTGCTCATATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1421  
Db 612 CTTTATCTCTCTGCTGCTCATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 671  
QY 1422 CAGCAAGGCTGAACGGAGGTGTGGAAGAGGAGCTGACGGTTCACCTCTCTGCTGCTGCTG 1481  
Db 672 GGAGAAGGCTTCTCGGAGGATGCTGGATGAATGACAGTCAGGCTTTCCTTCCCAAG 731  
QY 1482 GGACACTATCTCATCAACAACCGCCAGGCTCTAGAGAGCCACCCCGTACCAGGAGCC 1541  
Db 732 TGATTCTAGCATGTTCAACAATAACCGCTCTCTCA----- 769  
QY 1542 CCGGCTCTGCTGGAATCCGCCACTTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1601  
Db 769 ---CCTAGTGAACAGGCTCCAACTCGACTTACGATCGCATCTTT----- 811  
QY 1602 GCTCTCCAATCAGCCTACCGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1661  
Db 811 -----CCTCTGCGCC 821  
QY 1662 GGGCCCCCCCCACACCGCTGGGCCCAACCCACCAACACCCAGGCTACAGTGGGGACTA 1721  
Db 822 TGACTACCAGGAGCCATCCAGCTGATACGAACACTCCAGATTTGCTCCAGGGAGGA 881  
QY 1722 TATGAGCCTGAGAGCCAGCGCCCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1781  
Db 882 GGAGTCAGGCTGCGAGCGGTGTTGTGAAGCCAGTCCAGCCCTGAGGGGTGCTG 941  
QY 1782 CCATTATGCGAGGCTGACATTTACCTGCGAGGCGCTACCGGGGCAACACCTATGCT 1841  
Db 942 CCATATGAGAGGCTGACATAGTGAACCTCCAGAGGTGACAGGAGGCAACACATATCTC 1001  
QY 1842 TGTGCT 1898  
Db 1002 AGTGCT 1061  
QY 1899 TCGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1958  
Db 1062 CAGGAACCTCTAACTTTCAAGAGAGAGTGGGAGAGGACAGTGTGGGAGGTTCTATCT 1121  
QY 1959 GTGTGAGGTGCGACAGCCCTCAAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2018  
Db 1122 CTTGAGTGGAGGAGTGAAGATTTCAAGAGAGAGTGTGGCTGCTGCTGCTGCTGCTGCTGCT 1181  
QY 2019 GGGACACCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2078  
Db 1182 CAACAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1241  
QY 2079 GAATGATTTCTGAAAGAGGTGAAGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2138  
Db 1242 GAATGATTTCTTAAAGAGATTAAGATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1301  
QY 2139 GCT 2198  
Db 1302 TCTATGATGCT 1361  
QY 2199 CGGCGACCTCAACAGTTCCTCAGTCCGCCACCCAGCTGGAGGACAAAGGACCGCGGGGGC 2258  
Db 1362 TGGAGATCTCAATCAGTTCTTCTCCCGCACCGCCCTAATTCTTCTCTCCAGCG-- 1418  
QY 2259 CCCTGGGAGCGGAGGCTGCGCGAGGGGCCACCATCATGCTACCAATGCTGCTGCTGCTGCTGCT 2318

SQL	Sequence	272 BP;	52 A;	76 C;	72 G;	60 T;
	Query Match	6.8%;	Score 248;	DB 1;	Length 272;	
	Best Local Similarity	95.6%;	Pred. No. 4.8e-42;			
	Matches 259;	Conservative 0;	Mismatches 11;	Indels 1;	Gaps	
QY	2998	GATCCTCTCCACCCCTCCTCTAGCCATCCCTTGGGGAAGGTTGGGAGAAATATAGGATAG	3057			
Db	1	GATCCTCTCNACCCCTCCTCTAGCCATCCCTTGGGGAAGGTTGGGGAAGAAATATAGGATAG	60			
QY	3058	ACACTGGACATGGCCCATTTGGAGCACCTGGGCCCCACCTGGGACACACTGATTCCTCGGAGA	3117			
Db	61	ACACTGGACATGGCCCATTTGGAGCACACGCGCCCATCTGGACACACTGATTCCTCGGAGA	120			
QY	3118	GGTGGCTGGGCCCCCAGCTTCTCTCTCCCTGTACACACTGGAGCCCACTGGCTGAGAAT	3177			
Db	121	GGTGGCTGCNCCCCCAGCTTCTNTCTCCCTGTNACACACTGGACCCCACTGGCTGAGAAT	180			
QY	3178	CTGGGGGTGAGGAGACAAAGAGAGAGAAATGTTTCTTGTGCGCTGCTCTGTACTT	3237			
Db	181	CTGGGGGTGAGNGGACAAAGAGAGAGAAATNTTTCCTTGTGCGCTGCTCTGTACTN	240			
QY	3238	GTCTCAGCTT-GGGCTTCTTCTCCTCAT	3267			
Db	241	GTCTCAGCTTGGGNTTCTTCTCCTCCAT	271			

RESULT 13  
T51456  
ID T51456 standard; DNA; 2820 BP.  
AC T51456;  
DT 30-APR-1997 (first entry)

KW	kinase receptor activation; KIRA; receptor protein tyrosine kinase;
KW	rPK; signal transduction; autophosphorylation; gp.trkA;
KW	enzyme-linked immunosorbent assay; ELISA; ds.
OS	Chimeric herpes simplex virus;
OS	Chimeric Homo sapiens.
FS	Key
FS	Location/Qualifiers
FT	misc_rna
FT	5
FT	/*tag= a
FT	/note= "sp6 RNA start"
FT	65
FT	misc_rna
FT	/*tag= b
FT	/note= "begin gD from pchadIII"
FT	293..2740
FT	cds
FT	/*tag= c
FT	/product= gD.trkA fusion protein
FT	380..460
FT	/*tag= d
FT	/note= "trkA signal sequence"
FT	461..2737
FT	/*tag= e
FT	/note= "mature trkA"
FT	2778
FT	polya_signal
FT	/*tag= f
FT	/note= "SV40 early polyA"
PN	WO9514930-A1.
PD	01-JUN-1995.
PD	18-NOV-1994; U13329.
PR	23-NOV-1993; US-157563.
PR	20-DEC-1993; US-170558.
PR	05-AUG-1994; US-286305.
PI	(CETH) GENENTECH INC.
PI	Godowski PJ, Mark MR,
PI	WPI: 95-207042/27.
DR	P-PSDB: W11940.
PT	Measuring auto-phosphorylation of tyrosine kinase receptor - to
PT	Identify and characterise potential (ant)agonists of TKR
PS	Example 3; Fig12a-g; 148pp; English.
CC	CNA constructs (#51456-58) respectively code for gD.trkA (W11940),
CC	gD.trkB (W11941) and gD.trkC (W11942), comprising fusions between

CC herpes simplex virus glycoprotein D flag polypeptide and human  
CC receptor protein tyrosine kinases (rptks) trkA, B and C, which are  
CC involved in the interaction between neurotrophins and their target  
CC cells. The gp.trk constructs can be expressed e.g. in dpl2.CHO  
CC host cells and used in kinase receptor activation, enzyme-linked  
CC immunosorbent assays (KIRA ELISA) to measure autophosphorylation of  
CC rptk kinase domains. These assays are used to identify and  
CC characterise potential (ant)agonists of rptks, study ligand-receptor  
CC interactions, measure phosphatase activity and identify rptk ligands  
CC in clinical samples.  
SQ Sequence 2820 BP; 529 A; 867 C; 836 G; 588 T;

Query Match 5.4%; Score 195.2; DB 1; Length 2820;  
Best Local Similarity 59.6%; Pred. No. 4.8e-31;  
Matches 392; Conservative 0; Mismatches 248; Indels 18; Gaps 3;  
QY 1988 TCAGTCTTGAATTTCCCTTAAATGTCGTAGGGACACCCCTTGTGCTAGCTGTCAAGA 2047  
DB 1920 TCCTTCTGAGTGCCACACACCTCTCTGCTGAGCAGCAAGATGCTGTGCTGTCAAG- 1979  
QY 2048 TCTTACGGCCAGATGCCACCAAGAAATGCCAGGAATGATTTCTTGAAGAGGTGAAGATCA 2107  
DB 1979 --GCACCTGAAGGAGGCGTCCGAGAGTGTCTCGCAGGACTTCCAAGGTGAGCTGAGCTGC 2036  
QY 2108 TGTGAGGCTCAGGACACCAACATCATTCGGCTGCTGGGCTGTGTGTGTCAGGACGACC 2167  
DB 2037 TCACCATGCTGCAGCACCAGCAGCATCTGTCGCTTCTTCGGGCTGTGCACCGAGGCGCGCC 2096  
QY 2168 CCCTCTGATGATTAATGACTACATGAGAGAACGGGACCTCAACAGTTCTCTCAGTGGCC 2227  
DB 2097 CCCTGCTCATGTCTTTGATGATATGCGCAGCGGACCTCAACGGTCTCTCCGATCCC 2156  
QY 2228 ACCAGCTGGAGGACAAAGCAGCGGAGGCGCCCTTGGGACGGGACGCTCGCAGGCGGC 2287  
DB 2157 ATGGACCTGATGCCAA-----GCTGCTGGCTGGTGGGAGGATGTGCTCCAG 2204  
QY 2288 CCACCATCAGCTACCAATGCTGCTGATGTGGAGCCCGACAGATGCCCTCCGGATCGGCT 2347  
DB 2205 GCCCCTTGGGTGCGGAGGCTGTGCGCGTGGCTAGCCAGGTGCTGCGGGGATGGTGT 2264  
QY 2348 ATCTGGCCACACTCAACTTTTGTACATCGGACCTTGGCCACCGGAACCTGCTAGTGGGG 2407  
DB 2265 ACCTGGCGGTCTGATTTTGTGACCGGGACCTGGCCACAGCAACTGTCTAGTGGGCC 2324  
QY 2408 AAAATTTCAACATCAAAATCGCAGACTTTGGCATGAGCCGGAACCTCTATCTGGGACT 2467  
DB 2325 AGGGACTGGTGTCAAGATTGCTGATTTTGGCATGAGCAGGAGATATCTACAGCACCGACT 2384  
QY 2468 ATTACCGGTGTCAGGCGCGGACGTGCTGCCATCCCTGGATGGCTGGAGTGCATCC 2527  
DB 2385 ATTACCGGTGCGGAGGCGCGACCATGCTGCCCATTCGCTGGATGCCGCCCGAGAGCATCC 2444  
QY 2528 TCATGGGAAAGTTCACGACTCGAGTACGTGTGGGCTTTGGTGTGACCTGTGGGAGG 2587  
DB 2445 TGTAACCGTAAGTTCACACCGAGAGCGAGCTGTGGAGTTCGGCGTGGTGTCTGGGAGA 2504  
QY 2588 TGCTGATGCTGTGAGGCGCCAGCGCTTTGGGACGCTACCCAGCAGCAGGTCTATCGA 2645  
DB 2505 TCTTACCTACGGCAAG---CAGCCCTGGTACCAGCTCTCCCAACACGAGGCAATCGA 2559

## RESULT 14

ID V20445 standard; DNA; 2301 BP.  
AC V20445;  
DT 17-JUN-1998 (first entry)  
DE Human c-trk oncogene.  
KW Human; oncogene; proto-oncogene; neoplastic disease; anticancer;  
OS Homo sapiens.  
PN US5734039-A.  
PD 31-MAR-1998.

PF 15-SEP-1994; 306691.  
PR 15-SEP-1994; US-306691.  
PA (UYJE-) UNIV JEFFERSON THOMAS.  
PI Calabretta B, Skorski T;  
DR WPI; 98-229882/20.  
PT Anticancer composition comprising two anti-sense oligo:nucleotide(s)  
PT - targeting cytoplasmic and nuclear oncogene(s)  
PS Claim 1; Column 99-102; 92pp; English.  
CC The present invention represents an oncogene from the present invention.  
CC The present sequence describes a composition which comprises two  
CC antisense oligonucleotides. The first oligonucleotide is specific for a  
CC cytoplasmic oncogene or proto-oncogene selected from ras, raf, EGF-1,  
CC c-fms, c-ros, c-kit, c-met, c-trk, c-src, c-abl, bcr-abl, c-fgr and  
CC proto-oncogene selected from myc, jun, c-ets, c-fos, c-myb, B-myb,  
CC c-rel, c-vav, c-ski, c-spi, cyclin D1, PML/RAR alpha, AML1/MTG8,  
CC E2A/p1 and ALL-1/AF-4. The composition is used for treating cancer.  
CC The combination of antisense oligonucleotides has synergistically  
CC enhanced ability to inhibit growth of cancer cells.  
SQ Sequence 2301 BP; 545 A; 568 C; 785 G; 403 T;

Query Match 5.3%; Score 193.6; DB 1; Length 2301;  
Best Local Similarity 59.4%; Pred. No. 9.6e-31;  
Matches 391; Conservative 0; Mismatches 249; Indels 18; Gaps 3;  
QY 1988 TCAGTCTTGAATTTCCCTTAAATGTCGTAGGGACACCCCTTGTGCTAGCTGTCAAGA 2047  
DB 1275 TCCTTCTGAGTGCACACACCTCTCTGCTGAGCAGCAAGATGCTGTGCTGTCAAG- 1334  
QY 2048 TCTTACGGCCAGATGCCACCAAGAAATGCCAGGAATGATTTCTTGAAGAGGTGAAGATCA 2107  
DB 1334 --GCACCTGAAGGAGGCGTCCGAGAGTGTCTCGCAGGACTTCCAACGTGAGGCTGAGCTGC 1391  
QY 2108 TGTGAGGCTCAGGACCAACACATCATTCGGCTGCTGGGCTGTGTGTGTCAGGAGGACC 2167  
DB 1392 TCACCATGCTGAGCAGCAGCAGCATCTGCTGCGCTTCTTCGGCGTCTGCACCGAGGCGCGCC 1451  
QY 2168 CCCTCTGCTGATTAATGACTACTACTATGAGAGACGGGACCTCAACAGTTCCTCAGTGGCC 2227  
DB 1452 CCCTGCTCATGTTCTCGAGTATATGCGCAGCGGAGACCTCAACCGTCTCTCCGATCCC 1511  
QY 2228 ACAGCTGGAGGACAAAGCAGCGGCGGCGCCCTTGGGACGGGAGGCTGCGCAGGCGC 2287  
DB 1512 ATGGACCCGATGCCAA-----GCTGCTGGCTGGTGGGAGGATGTGGCTCCAG 1559  
QY 2288 GCACCATCAGCTACCAATGCTGCTGCATGTGGCAGCCCGAGATCGCTCCGCGATGCGCT 2347  
DB 1560 GCCCCTGGGTCTGGGCGAGCTGCTGGCGTGGCTAGCCAGTCTGCGGGGATGGTGT 1619  
QY 2348 ATCTGGCCACACTCAACTTTGTACATCGGGACCTGGCCACGCGGAACCTGCCCTAGTGGGG 2407  
DB 1620 ACCTGGCGGTCTGCAATTTTGTGACCGGGACCTGGCCACAGCAACTGTCTAGTGGGCC 1679  
QY 2408 AAAATTTCAACATCAAAATCGCAGACTTTGGCATGAGCCGGAACCTCTATGCTGGGACT 2467  
DB 1680 AGGAGTGTGTGAAGATTGCTGATTTTGGCATGAGCAGGAGATATCTACACACCGACT 1739  
QY 2468 ATTACCGTGTGAGGCGCGGAGTGTGCTGCCCATCGCTGGATGGCCTGGGAGTGCATCC 2527  
DB 1740 ATTACCGTGTGGAGGCGCGACCATGCTGCCCATTCGCTGGATGCGGCCCGAGAGCATCC 1799  
QY 2528 TCATGGGAAAGTTCACGACTGCGAGTGTGGGCTTTGGTGTGACCTGTGGGAGG 2587  
DB 1800 TGTACCGTAAGTTCACCGAGAGGAGCTGTGGAGTTCGGCGTGGTGTCTCTGGGAGA 1859  
QY 2588 TGCTGATGCTGTAGGCGCCAGCGCTTTGGGACGCTCACCAGCAGGAGGAGTCAATCGA 2645  
DB 1860 TCTTACCTACGGCAAG---CAGCCCTGGTACCAGCTCTCCCAACACGAGGCAATCGA 1914

## RESULT 15

Q99277



0y	1	MGPEALSS-LLLLLLVASGDADMKGHFDPACRYALGNQDRTIPDSISASSWSHSDSTAA	59
0y	60	RHSRLSSDGDGAWCPAGSVFFKEEYLVQDLRLHLVALVGTGRHAGGLGKEFSRSYR	119
0y	61	RHSRLSSDGDGAWCPAGVPVFEKEEYLVQDLRLHLVALVGTGRHAGGLGKEFSRSYR	120
0y	120	LYSRDGRRMWKKRWGQEVISGNEDPEGVVLKDLGPPMWARLVRFYPRADRVMSVCLR	179
0y	121	LYSRDGRRMWKKRWGQEVISGNEDPEGVVLKDLGPPMWARLVRFYPRADRVMSVCLR	180
0y	180	VELYGCILWRDGLLSYTPAVGQTMVLSA-VYINDSTYDGHVTGGLQYGLGLADGWVGL	238

```
Db 181 VELYGLWRDGLLSYAPVGTMTOLSEVMVHLNDSTVDTAGGLQYGGGLQGLADGVVGL 240
Qy 239 DDFRKSQELRWPGYDYVGNVSHSFGYVEMEFEDRLRAFAQMVHCHNMHTLGLARLP 298
Db 241 DDFRKSQELRWPGYDYVGNVSHSFGYVEMEFEDRLRTFTQMVHCHNMHTLGLARLP 300
Qy 299 GGVSECRFRGPAWMEGEPVHNGNGLGDPRAVSVPLGGRRVAFLOCRFLFAGPWLL 358
Db 301 GGVSECRFRGPAWMEGEPVHNGNGLGDPRAVSVPLGGRRVAFLOCRFLFAGPWLP 360
Qy 359 FSEISFISDVVNNSSPALGCTEPAPWMPGPPPTNFSSLELEPRGQOPVAKPEGSPAT 418
Db 361 FSEISFISDVVNNSS---DTEFPAPWMPGPPPTNFSSLELEPRGQOPVAKPEGSPAT 416
Qy 419 LIGLVAIILLLLIIALMLWRLHWRLLSKAERRVLEELTVHLSVPGDTIILNNRPGP 478
Db 417 LIGLVAIILLLLIIALMLWRLHWRLLSKADSRVLEELTVHLSVPGDTIILNNRPGP 476
Qy 479 REPPYQEPFRGPNPHSAPCVNGSALLSNPAYRLLLATYARPPRGPGPPTPAWAKPT 538
Db 477 REPPYQEPFRGPNPHSAPCVNGSALLSNPAYRLLLATYARPPRGPGPTPAWDKPT 536
Qy 539 NTQAYSGDYMEPEKPGAPLPPPPQNSVPHYAEADIVTLQVGTGGNTYAYPALPPGAVGD 598
Db 537 NTQAYSGDYMEPEKPGAPLPPPPQNSVPHYAEADIVTLQVGTGGNTYAYPALPPGAVGD 596
Qy 599 GPPRVDFPRSLRKEKLGEGFGEVHCEVDSPQDLVSLDFPLNVRKHGHPLLVAVKILR 658
Db 597 GPPRVDFPRSLRKEKLGEGFGEVHCEVDSPQDLVSLDFPLNVRKHGHPLLVAVKILR 656
Qy 659 PDATKNARNDFLKVKIMSLRKLDPNIIRLLGVCVQDDPLCMITDYMGNDLQFLSAHQL 718
Db 657 PDATKNARNDFLKVKIMSLRKLDPNIIRLLGVCVQDDPLCMITDYMGNDLQFLSAHQL 716
Qy 719 EDKAEAGPQGOAQAQPTISYPMLLHVAQIASGMRYLATLNFVHRDLATRNCLVGENF 778
Db 717 ENKATQGLSGTSDQDQPTISYPMLLHVAQIASGMRYLATLNFVHRDLATRNCLVGENF 776
Qy 779 TIKIADFGMSRLNLAGDYRVOGRAVPLIRWMAWECILMGKFTTASDVAFGVTLEVLML 838
Db 777 TIKIADFGMSRLNLAGDYRVOGRAVPLIRWMAWECILMGKFTTASDVAFGVTLEVLML 836
Qy 839 LCRAQPGQLTDEQVIENAGFEFFRQGRQVLSRPPACPOGLVELMLRCWRSRSEQRPPF 898
Db 837 LCRAQPGQLTDEQVIENAGFEFFRQGRQVLSRPPACPOGLVELMLRCWRSRSEQRPPF 896
Qy 899 SOLHRELAEDALNTV 913
Db 897 AQLHRELAEDALNTV 911
RESULT 2
ID Q64108 PRELIMINARY; PRT; 220 AA.
AC Q64108;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE RECEPTOR TYROSINE KINASE (FRAGMENT).
GN PTK-3L.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95320273.
RA SAKUMA S., SAYA H., IJICHI A., TOFILON P.J.;
RT "Radiation induction of the receptor tyrosine kinase gene Ptk-3 in
RL normal rat astrocytes."
DR Radiat. Res. 143:1-7(1995).
DR EMBL; S77585; AAB34728.1; -.
ET NON_TER 1
SQ SEQUENCE 220 AA; 23670 MW; 807900DF CRC32;
```

```
Query Match 23.5%; Score 1154; DB 11; Length 220;
Best Local Similarity 98.2%; Pred. No. 8.5e-90;
Matches 216; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 404 GQOPVAKPEGSPATLIGCLVAIILLIIALMLWRLHWRLLSKAERRVLEELTVHL 463
Db 1 GQOPVAKAEGSPATLIGCLVAIILLIIALMLWRLHWRLLSKAERRVLEELTVHL 60
Qy 464 SVPGDTIILNNRPGPREPPYQEPFRGPNPHSAPCVNGSALLSNPAYRLLLATYARP 523
Db 61 SVPGDTIILNNRPGPREPPYQEPFRGPNPHSAPCVNGSALLSNPAYRLLLATYARP 120
Qy 524 PRGPGPPTPAWAKPTNTQAYSGDYMEPEKPGAPLPPPPQNSVPHYAEADIVTLQVGTGG 583
Db 121 PRGPGPPTPAWAKPTNTQAYSGDYMEPEKPGAPLPPPPQNSVPHYAEADIVTLQVGTGG 180
Qy 584 NTYAVPALPPGAVGDGPPRVDFPRSLRKEKLGEGFGE 623
Db 181 NTYAVPALPPGAVGDGPPRVDFPRSLRKEKLGEGFGE 220
RESULT 3
ID Q19354 PRELIMINARY; PRT; 791 AA.
AC Q19354;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE SIMILAR TO TYROSINE KINASE. NCBI GI: 1118035.
GN FL1D5.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 94150718.
RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KESHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SALLDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RN elegans."
RN [2]
RP SEQUENCE FROM N.A.
RA NHAN M.;
RN [3]
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.
RA WATERSTON R.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U41532; AAB3276.1; -.
DR PFAM; PF00754; F5_F8_type-C; 1.
DR PFAM; PF00069; pkinase; 1.
SQ SEQUENCE 791 AA; 89990 MW; 991C9D67 CRC32;
```

```
Query Match 19.7%; Score 968; DB 5; Length 791;
Best Local Similarity 28.4%; Pred. No. 2.8e-73;
Matches 253; Conservative 152; Mismatches 309; Indels 178; Gaps 31;
Qy 46 DISASSSSSDSTAAR-HSRLESSDGDGAWCPAGSVFPKKEEYLVQDLQRLHLVALVGTQ 104
Db 34 DLNMLGKWFNGVYFIRTAHQESGSGAWCPKQNSLKSKEWLQISFSDVITVSVETQ 93
```

RT	"Radiation induction of the receptor tyrosine kinase gene Ptk-3 in normal rat astrocytes."
RL	Radiat. Res. 143:1-7(1995).
DR	EMBL; S77556; AAB34729.1; -.
FT	NON_TER
SQ	SEQUENCE 183 AA; 19717 MW; OBIAD57B CRC32;
Query Match	18.9%; Score 924.5; DB 11; Length 183;
Best Local Similarity	81.4%; Pred. No. 1.6e-70;
Matches 179; Conservative	0; Mismatches 4; Indels 37; Gaps
QY	404 GOOPVAKPEGSPTAILGCLVAIIILLIALLMLWLHWRLLSKAERRVLEELTVHL 463
Db	1 GOOPVAKAEGSPTAILGCLVAIIILLIALLMLWLHWRLLSKAERRVLEELTVHL 60
QY	464 SVPGDTILINNRCPREPPPYQEPFRGNPHPSAPCVNGSALLLSNPAYRULLLATYARP 533
Db	61 SVPGDTILINNRCPREPPPYQEPFRGRPTHSACVPNGSNAC ----- 104
QY	524 PRGCGPPTPAWAKTNTQAYSGDYMEPEKGPGLPLPPPQNPSVPHYAEADIVTLQGVTGG 583
Db	104 -----SGDYMEPEKGPGLPLPPPQNPSVPHYAEADIVTLQGVTGG 143
QY	584 NTYAVPALPGCAVGDPGRVDPRSRRLKEKLGEGQGE 623
Db	144 NTYAVPALPGCAVGDPGRVDPRSRRLKEKLGEGQGE 183
RESULT	5
O18433	PRELIMINARY; PRT; 700 AA.
ID	O18433
AC	O18433;
DC	01-JAN-1998 (TrEMBLrel. 05, Created)
DT	01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT	01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE	RECEPTOR TYROSINE KINASE
OS	Geodia cydonium (Sponge)
OC	Eukaryota; Metazoa; Porifera; Demospongiae; Tetractinomorpha;
CC	Astrophorida; Geodiidae; Geodia.
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE; 97224609.
RA	GAMULIN V., SKORKHOOD A., MUELLER M., SCHAECKE H., MUELLER W.E.G.;
RT	"Experimental indication in favor of the introns-late theory: the receptor tyrosine kinase gene from the sponge Geodia cydonium.";
RL	J. Mol. Evol. 44:242-252(1997).
DR	EMBL; X94128; CAA63848.1; -.
DR	PFAM; PF00047; ig; 1.
DR	PFAM; PF00069; pkinase; 1.
SQ	SEQUENCE 700 AA; 77850 MW; 2F0C4179 CRC32;
Query Match	14.3%; Score 702; DB 5; Length 700;
Best Local Similarity	32.8%; Pred. No. 6.8e-51;
Matches 193; Conservative	85; Mismatches 191; Indels 120; Gaps
QY	413 GSPTAILGCLVAIIILLIALL--MWRLHWRR-----ILSK 449
Db	140 GNSGVIAAGVLITLIILLIALLIFVFVVGYRRRGKFDLGSCRELSCGSCVCVPLLA 199
QY	450 AERRVLEELTVHLSVPGDPTILINNPPGPREP----PPYOEPFRGN--PPHSAPCPVN 502
Db	200 LAGVKLPTRHRLENKNGTRLRNENHTADNTNIYSVWORPKKKIKSGPPLPP ---- 256
QY	503 GSALLSNPAYRLLLATYARPPGPGPTPAWAKPTNQAYSGDYMEPEKGPGLPLPPP 562
Db	256 ---LTLETNELMSIDEELSPIOEKPT--RRNTGLSTYSQSQTIPKLAKUTLKRFK 311
QY	563 QNSVPHYAEADIVTLQ-GVTGGNT-YAVPALP----- 593
Db	312 MKENPIQSVDDVILELVLEQLQNTALYPALKFNSTNSASFTDDLADSPITYSVAINPSMF 371

QY 593 ---PGAVGD-----GP-----PRVDFPRSLRFEKELGSGQGEVHLCVD- 631  
Db 372 TKRSITIGNDLHPYGIYARPIKQKMRQPLNSVDNIREVKOIGVGQFVLAEMTG 431  
QY 631 -SPQDLVSLDFPLNVRKH-----PLLVAVKILRPDATKNARNDFLKEVKIMSLKDPNI 684  
Db 432 LSGSNVASLP-----KGSMDAGVALVAVKILKPDVSDVSRFQDFKEIKFMSQLQHDSI 485  
QY 685 IRLGVCVQDDPLCMITDYMENGLDNQFLSAHQLEDAAGAGDGOAAGQPTISYPMLL 744  
Db 486 VOLLAVCTHSHKPFVMEYMEENGLDNQFLQYQWVD-----DDSAIYSNQIPPTSTLL 537  
QY 745 HVAQIAGSMRYLATLNFVHRDLATRNCLVGENFTIKIADFGMSRNLVYAGDYRVQGRAV 804  
Db 538 YMAVQIAGSMRYLSSNLNVHRDLATRNCLVGENFTIKISDFGMSRNLVYRVYVRGRAM 597  
QY 805 LPIRMAMECILMGKFTTASDVWAGVTLVLEVMCLRAQPFQGLTDEQVIENAGEFFRDQ 864  
Db 598 LPIRMAYTES-FYGRFSKSDAWAYGVTVWEIYTLGKKQPYEELDDQHMQDA---IRGT 653  
QY 865 GRQVLSRPPACPGGLYELMLRCRSESEQRPPFSQHLRFLAEDALNTV 913  
Db 654 GRRI-MGRPECPQAVYEVLLRCWEYAADRAATFKEIH-----DSLNL 696  
RESULT 6  
QYH44  
ID QYH44 PRELIMINARY; PRT; 821 AA.  
AC QYH44;  
DT 01-MAY-1999 (TRENBLrel. 10, Created)  
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
DE NEUROTROPHIN RECEPTOR B XTRKB-ALPHA.  
GN XTRKB.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Amphibia; Batrachia; Anura;  
OC Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA ISLAM N., GAGNON F., MOSS T.;  
RT "Catalytic and non-catalytic forms of the neurotrophin receptor xtrkb  
RT mRNA are expressed in a pseudo-segmental manner within the early  
RT Xenopus central nervous system."  
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U39670; AAD00001.1;  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS00239; RECEPTOR\_TYR\_KIN\_II; 1.  
KW Receptor.  
SQ SEQUENCE 821 AA; 92311 MW; DSCBCD27 CRC32;

Query Match 13.2%; Score 646; DB 13; Length 821;  
Best Local Similarity 45.6%; Pred. No. 4.6e-46;  
Matches 139; Conservative 48; Mismatches 80; Indels 38; Gaps 8;  
QY 607 RSLRFEKELGSGQGEVHLC-----VDSQDVLVSLDFPLNVRKHGHPLLVAVKILRPDATK 663  
Db 534 RHNVILKRELGEAGFKVFLAECYNLYPEQD-----KILVAVKTLK-DASD 578  
QY 664 NARNDFLEKVKIMSLRKDPNIIRLLGVCVQDDPLCMITDYMENGLDNQFLSAHQLEDAK 723  
Db 579 NARKDFHREAELLTLQHEHIVKFGVCGVEGDLIMVFYMKHGDNLKFLRAH----- 632  
QY 724 EGAPGDGQAAG---PTISYPMLLHVAQAISGMRYLATLNFVHRDLATRNCLVGENFTI 780  
Db 632 --GPDVLAEGNRPAELTQSOMLHIAQIAGWYLASQHFVHRDLATRNCLVGENLLV 689  
QY 781 KIADFGMSRNLVYAGDYRVQGRAVLPIRMWAMECILMGKFTTASDVWAGVTLVLEVMCL 840  
Db 690 KIGDFGMSRDVYTDYRVGGHTMLPIRMWPPESIMYRKFTTESDVWSLGVVLWEIFTYG 749  
QY 841 RAQPFQGLTDEQVIENAGEFFRDQGRQVYLSRPPACPGGLYELMLRCRSESEQRPPFSQ 900

Db 750 K-QPWYQLSNNEVIECI-----TQGR--VLQRPRTCPKEVYDLMLGCGWQREPHMRLNIKE 801  
QY 901 LHRFL 905  
Db 802 IHSLL 806  
RESULT 7  
QYH43  
ID QYH43 PRELIMINARY; PRT; 811 AA.  
AC QYH43;  
DT 01-MAY-1999 (TRENBLrel. 10, Created)  
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)  
DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
DE NEUROTROPHIN RECEPTOR B XTRKB-ALPHA.  
GN XTRKB.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Amphibia; Batrachia; Anura;  
OC Mesobatrachia; Pipoidae; Pipidae; Xenopodinae; Xenopus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA ISLAM N., GAGNON F., MOSS T.;  
RT "Catalytic and non-catalytic forms of the neurotrophin receptor xtrkb  
RT mRNA are expressed in a pseudo-segmental manner within the early  
RT Xenopus central nervous system."  
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U39671; AAD00002.1;  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS00239; RECEPTOR\_TYR\_KIN\_II; 1.  
KW Receptor.  
SQ SEQUENCE 811 AA; 91249 MW; B267BDE7 CRC32;  
Query Match 13.2%; Score 646; DB 13; Length 811;  
Best Local Similarity 45.6%; Pred. No. 4.5e-46;  
Matches 139; Conservative 48; Mismatches 80; Indels 38; Gaps 8;  
QY 607 RSLRFEKELGSGQGEVHLC-----VDSQDVLVSLDFPLNVRKHGHPLLVAVKILRPDATK 663  
Db 524 RHNVILKRELGEAGFKVFLAECYNLYPEQD-----KILVAVKTLK-DASD 568  
QY 664 NARNDFLEKVKIMSLRKDPNIIRLLGVCVQDDPLCMITDYMENGLDNQFLSAHQLEDAK 723  
Db 569 NARKDFHREAELLTLQHEHIVKFGVCGVEGDLIMVFYMKHGDNLKFLRAH----- 622  
QY 724 EGAPGDGQAAG---PTISYPMLLHVAQAISGMRYLATLNFVHRDLATRNCLVGENFTI 780  
Db 622 --GPDVLAEGNRPAELTQSOMLHIAQIAGWYLASQHFVHRDLATRNCLVGENLLV 679  
QY 781 KIADFGMSRNLVYAGDYRVQGRAVLPIRMWAMECILMGKFTTASDVWAGVTLVLEVMCL 840  
Db 680 KIGDFGMSRDVYTDYRVGGHTMLPIRMWPPESIMYRKFTTESDVWSLGVVLWEIFTYG 739  
QY 841 RAQPFQGLTDEQVIENAGEFFRDQGRQVYLSRPPACPGGLYELMLRCRSESEQRPPFSQ 900  
Db 740 K-QPWYQLSNNEVIECI-----TQGR--VLQRPRTCPKEVYDLMLGCGWQREPHMRLNIKE 791  
QY 901 LHRFL 905  
Db 792 IHSLL 796  
RESULT 8  
QYH43  
ID QYH43 PRELIMINARY; PRT; 790 AA.  
AC QYH43;  
DT 01-NOV-1996 (TRENBLrel. 01, Created)  
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
DE TROPOMYOSIN RECEPTOR KINASE.  
GN CTRKA.



OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Aves;  
 OC Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA SCHROEPPEL A., V SCHACK D., DECHANT G., BARDE Y.A.;  
 RL Mol. Cell. Neurosci. 6:0-0(0).  
 DR EMBL; U43396; AAA85289.1; -.  
 DR PFAM; PF00560; LRR; 2.  
 DR PFAM; PF00069; pkinase; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS00239; RECEPTOR\_TYR\_KIN\_II; 1.  
 SQ SEQUENCE 790 AA; 88669 MW; B5F7768C CRC32;

Query Match 13.0%; Score 636.5; DB 13; Length 790;  
 Best Local Similarity 30.8%; Pred. No. 2.7e-45;  
 Matches 192; Conservative 78; Mismatches 189; Indels 165; Gaps 22;

QY 356 WLLFSEISFISDVVNSPALGGTFPPAPWPPGPPPTNFSSLELEPRGQQPVAKPEGSP 415  
 DB 237 WELVLEINNISSLNK-----DLTCRAENAVGLAEDS- 270  
 QY 416 TALLICGLVAIIILLIILALMLRLHWR-----RLLSKAERRVLEELTVHLSV--- 466  
 DB 270 -VMLNVTFFPVILLSEAIPOHFWCIPFSDVSNPTPRILWLFNGMLPEGPYIHRIVEY 328  
 QY 466 -PDGTILN----NRGPPREPPYQPRGRGNPHPSAPCPVNGS-ALLSNP---AYRL 516  
 DB 329 EPNSTVLHGCLQNR-----PTH-----VNNGNTLVVQNPGLGRATRSI 367  
 QY 517 LAYARPP--RGGPPTPAWAKPTNTQAYSGDYMEPEKFGAPLLPPPPQNSVPHYAEADI 574  
 DB 368 QGRFMDNPFSPPEEIPVISISPLGRNSS-----LEGPVETADEHTFGVSV 414  
 QY 575 VTLQGV-----TGNTYAYPA-----LPPGAVGGDP 600  
 DB 415 AVAVALFASLFLSVMLIALNKGHRKFGINRSVLAQEDDLAMSLHFMNLGSSPVSSTE 474  
 QY 601 PRVD-----FPRSLRLEKEKGQFGEVHLCEVDS--PDLYSL 638  
 DB 475 SKLDGLKSNFIENPOYFCNACVHHVQRDIVLKLWELGEGAFKGVLAECSHLLPQ----- 531  
 QY 639 DFPLNVRKGHPLLVAVKILRPDATKNARNDFLKEYKIMSLRKPDIIRLLGVCVDDPLC 698  
 DB 531 -----EKLTVAVKALK-EVTESARLDQFREAELLTVLQHEHIVKFGVCTGEPLI 580  
 QY 699 MITDYMENGLAOLFSAHOLEDKAAEGAPGCGQAAGQPTISYPMLLHVAAQIASGMRYLA 758  
 DB 581 MVEFYMKHGLDARFLRSHGPDAKILD--OGOGQPCGQLTSLH--MLQIATQIASGMVYLA 636  
 QY 759 TLNFVHRDLATRNCLVGENFTIKIADFGMSRLYAGDYRVQGRAVLPIRMAWECILMG 818  
 DB 637 SLHFVHRDLATRNCLVGLDVLKVGDFGMSRDIYSTDYRVGGRMLPIRMPESILYR 696  
 QY 819 KTTASDVWAFQVTLWEVLMCRAPFGQLTDEQVIENAGEFFRDQGRQVLYSRPPACQ 878  
 DB 697 KFTTESDIWSFGVWLWEIFTYKQ-QPWYQLSNTAEICI-----TQGRE--LERPRTCP 748  
 QY 879 GLVELMCRWSRESEORPFSOLH 902  
 DB 749 EYVDIMQSCWQREPOORIQDIH 772

RESULT 9  
 ID 075682 PRELIMINARY; PRT; 839 AA.  
 AC 075682;  
 DT 01-NOV-1998 (Tremblrel. 08, Created)  
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)  
 DT 01-MAY-1999 (Tremblrel. 10, Last annotation update)

DE TRKC PROTEIN.  
 GN TRKC.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
 OC Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA ICHASO N., RODRIGUEZ R., MARTIN-ZANCA D., GONZALEZ-SARMIENTO R.;  
 RL "Genomic characterization of the human TRKC gene";  
 RT Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ224521; CAAL2029.1; -.  
 DR EMBL; AJ224522; CAAL2029.1; JOINED.  
 DR EMBL; AJ224523; CAAL2029.1; JOINED.  
 DR EMBL; AJ224524; CAAL2029.1; JOINED.  
 DR EMBL; AJ224525; CAAL2029.1; JOINED.  
 DR EMBL; AJ224526; CAAL2029.1; JOINED.  
 DR EMBL; AJ224527; CAAL2029.1; JOINED.  
 DR EMBL; AJ224528; CAAL2029.1; JOINED.  
 DR EMBL; AJ224529; CAAL2029.1; JOINED.  
 DR EMBL; AJ224530; CAAL2029.1; JOINED.  
 DR EMBL; AJ224531; CAAL2029.1; JOINED.  
 DR EMBL; AJ224532; CAAL2029.1; JOINED.  
 DR EMBL; AJ224533; CAAL2029.1; JOINED.  
 DR EMBL; AJ224534; CAAL2029.1; JOINED.  
 DR EMBL; AJ224535; CAAL2029.1; JOINED.  
 DR PFAM; PF00560; LRR; 2.  
 DR PFAM; PF01463; LRRCT; 1.  
 DR PFAM; PF01462; LRRNT; 1.  
 DR PFAM; PF00089; pkinase; 1.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
 DR PROSITE; PS00239; RECEPTOR\_TYR\_KIN\_II; 1.  
 SQ SEQUENCE 839 AA; 94428 MW; 986D7E71 CRC32;

Query Match 12.9%; Score 633.5; DB 4; Length 839;  
 Best Local Similarity 32.0%; Pred. No. 5.4e-45;  
 Matches 174; Conservative 66; Mismatches 132; Indels 171; Gaps 19;

QY 490 RGNPPHSPACVPNG-----SALLSNPA-----YRLLATYAR 522  
 DB 326 RGNPPTLHLWHNGQPLRESKIIHVEYQGEISEGCLLFNKRPTHYNNGYTLI-----AK 381  
 QY 523 PPRGPGPTPAWAKPTNTQAYSGDYME---PEKPGAPL---PPPPQNSVPHYAEADIV 575  
 DB 382 NPLG-----TANQINGHFLKEPPESTDFILFDEVPSTPIVTHKPEDIF 430  
 QY 576 TLOGVTGNTYAVDAL-----P 592  
 DB 431 GVSIAVGLAAFAFACVLLVFLVMINKYGRSKFGMKGPVAVISGEEDSASPLHHNHGITT 490  
 QY 593 PGAVGGDP-----PRVDFP-----RSRLRFEKELGEGQFG 622  
 DB 491 PSSLDAGPTIVTGMTRIPVIEPQVROGHNCHKPDTYVQHKRDRDIVLKRLEGEGAFG 550  
 QY 623 EVHLCVDSQDLVSLDFPLNVRKHPLVAVKILRPDATKNARNDFLKEVKIMSLKDP 682  
 DB 551 KVFLAEC-----YNLSPTK-DKMLVAVKALK-DPTLAARKDFQREALLTNLQHE 598  
 QY 683 NITRLGVCVDDPLCMITDYMENGLNOFLSAHOLEDKAAEGAP-----DQAAAQGP- 737  
 DB 599 HIVKFGVCGDGLIMVFEYMKHGDNLKFLRAH-----GPDAMILVDGQPRQAKG 649  
 QY 737 TISVPMLLHVAAQIASGMRYLATLNFVHRDLATRNCLVGENFTIKIADFGMSRLYAGDY 796  
 DB 650 ELGLSQMLHIASQIASGMRYLASQHEVHRDLATRNCLVGENFTIKIADFGMSRDIYSTDY 709  
 QY 797 YR-----VOGRAVLPIRMAWECILMGKFTTASDVWAFQVTLWEVLMCR 842  
 DB 710 YRLNFPNGNDFCIWCEVGGHTMLPIRMPESIMYRKFTTESDVSFGVILMEITYGK- 769  
 QY 843 QPFGQLTDEQVIENAGEFFRDQGRQVLYSRPPACQGLYELMCRWSRESEORPFSOLH 902  
 DB 843 QPFGQLTDEQVIENAGEFFRDQGRQVLYSRPPACQGLYELMCRWSRESEORPFSOLH 902

Db 769 QPWFLSNTVEICI-----TQGR--VLERPRVCPKEVDYVLMGCMQREPQORLNKEIY 821

QY 903 RFL 905

Db 822 KIL 824

RESULT 10

Q27656 PRELIMINARY: PRT: 699 AA.

AC Q27656;

DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)

DE SPONGE RECEPTOR TYROSINE KINASE (EC 2.7.1.112)

DE (PROTEIN-TYROSINE KINASE) (TYROSYLPROTEIN KINASE) (PROTEIN KINASE

DE (TYROSINE)) (HYDROXYARYL-PROTEIN KINASE).

GN GCTK 2.

OS Geodia cydonium (Sponge).

OC Eukaryota; Metazoa; Porifera; Demospongiae; Tetractinomorpha;

OC Astrophorida; Geodiidae; Geodia.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 95142935.

RA GAMULIN V., RINKEVICH B., SCHAECKE H., KRUSE M., MUELLER I.M.,

RA WERNER E.G.;

RT "Cell adhesion receptors and nuclear receptors are highly conserved

RT from the lowest metazoa (marine sponges) to vertebrates.";

RL Biol. Chem. Hoppe-Seyler 375:583-588(1994).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE; 95251882.

RA SCHAECKE H., RINKEVICH B., GAMULIN V., MUELLER I.M., MUELLER W.E.G.;

RT "Immunoglobulin-like domain is present in the extracellular part of

RT the receptor tyrosine kinase from the marine sponge Geodia

RT cydonium.";

RL J. Mol. Evol. 44:242-252(1997).

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE; 97224609.

RA GAMULIN V., SKOROKHOD A., MUELLER M., SCHAECKE H., MUELLER W.E.G.;

RT "Experimental indication in favor of the introns-late theory: the

RT receptor tyrosine kinase gene from the sponge Geodia cydonium.";

RL J. Mol. Evol. 44:242-252(1997).

CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP + PROTEIN

CC TYROSINE PHOSPHATE.

DR EMBL; X77528; CAA54563.1; -

DR PFAM; PF00047; 1g; 1.

DR PFAM; PF00069; pkinase; 1.

KW Transferase.

SQ SEQUENCE 699 AA; 77477 MW; F65E4B75 CRC32;

Query Match 12.9%; Score 633.5; DB 5; Length 699;

Best Local Similarity 32.0%; Pred. No. 4.1e-45;

Matches 194; Conservative 81; Mismatches 175; Indels 157; Gaps 25;

QY 413 GSTAILGLCVAILLLLIIIAL--MLWRLHWRLLSKAERRVLEELTVHLS----- 465

Db 140 GSNVGVIAGVLTLLIIILIIILFVFWCYRR-----RGKLDGSCRLSCGSCSC 193

QY 465 VPGDTILINRRPGPREPPYPRGRNPPHSPAPVPGNSALLS-----NP 511

Db 194 VP---LLAALKGVKLPTRHRENLDK-----NGTRLRLNERNHIADTNTIYSV 238

QY 512 AYRLLIATYARPRGCGPP-----TPAWAKPT--NT--QAYSGDY 547

Db 239 VQPLKIKISKSP--PLPPLTLTETELNETLSIDEKEULSPQIEKPTRRRTGLSTISOG 296

QY 548 MEPEKPGAPLLPPPPQNSVPHAEADIVTLQ-GVTGGNT-YAVPALP----- 593

Db 297 TIPKLAKLTKLRFKKNENPIYQSADELELELELQVDNTLYALPSKPNSTRNSASFTDDL 356

QY 593 -----PGAVGD-----GP-----PRVDFPRSRLRFEKEL 616

Db 357 ASDPIYSVAINPSMFTKRSSTIGNDLHPYGIYARPIKQMRQPLNVSVDNREVQI 416

QY 617 GEGQGEVHLCEVD--SPQDLYSLDFPLNVRKGH-----PLLVAVKILRPDANKARND 669

Db 417 GVGQFGAVVLAEMTGLSGSNVASLP-----KGSMAAGVALVAVKLUKPDVSDVLSQSF 470

QY 670 LKEVKIMSRKDPNIIRILIGVCVQDDPLCMITDYMENGDLNQFSAHQLEDKAEGAPGD 729

Db 471 DKEIKFMSQLQHSIVQLLAICTHSHKHPFIYMEYMGDLNQFLQKYQWVD-----D 522

QY 730 GOAAGCTTISPMILHVAQAIAAGMYRIATLNFVHRDLATRNCLNGENFTTKIADFGMSR 789

Db 523 DSALSNQNPSTLLYMAVQIASGMVYLSLNVYHRDLATRNCLVGSNFRKISDFGMR 582

QY 790 NLYAGDYRVQGRAVLPPIRMMAWECILMGKFTTASDVMAFGVTLWEVLMLCRAQFPGOLT 849

Db 583 NLYERYVYRVGRAMLPIRMWATES-FYGRFSEKSDAWAIGVTWVEIITLGGKKQYEEEL 641

QY 850 DEQVIENAGEFFRDGQGVYLSRP---PACPOGLYELMLRCWSRESEQRPPEQLHRELA 906

Db 642 DQDMIQDA---IRGTGRRI-MGRPRGVACVGRG---ATRCWVYAAADRATFEKI-- 690

QY 907 EDALNTV 913

Db 690 -DSLNI 695

RESULT 11

Q15655 PRELIMINARY: PRT: 503 AA.

ID Q15655;

AC Q15655;

DT 01-NOV-1996 (TREMBlrel. 01, Created)

DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)

DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)

DE 55 KD PROTEIN.

GN TRK-T1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

OC Eutheria; Primates; Catarrhini; Hominiidae; Homo.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 92195050.

RA GRECO A., PIROTTI M.A., BONGARZONE I., PAGLIARDINI S., LANZI C.,

RA DELLA PORTA G.;

RT "TRK-T1 is a novel oncogene formed by the fusion of TPR and TRK genes

RT in human papillary thyroid carcinomas.";

RL Oncogene 7:237-242(1992).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE; 92175499.

RA HILL K., BOONE C., GOEBL M., PUCCIA R., SDICU A.M., BUSSEY H.;

RT "Yeast KRE2 defines a new gene family encoding probable secretory

RT proteins, and is required for the correct N-glycosylation of

RT proteins.";

RL Genetics 130:273-283(1992).

DR EMBL; X62947; CAA44719.1; -

DR PFAM; PF00069; pkinase; 1.

DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.

DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.

DR PROSITE; PS00239; RECEPTOR\_TYR\_KIN\_II; 1.

SQ SEQUENCE 503 AA; 58174 MW; E2D051E1 CRC32;

Query Match 12.8%; Score 628; DB 4; Length 503;

Best Local Similarity 45.5%; Pred. No. 7.5e-45;

Matches 136; Conservative 49; Mismatches 84; Indels 30; Gaps 8;

QY 607 RSLRFRKEKLGEGQFGEVHLCEVDS---PDLYSLDFPLNVRKGHPLLVAVKILRPDANK 663

Db 214 RRDVLKWLKEGAFGKVFLEAECNLLPEQD-----KMLVAVKALK-EASE 258

```

QY 664 NARNDLKEVKIMSLKDPNIIRLLGVCVQDDPLCMITDYMNGDNLQFSLAHLQEDKAA 723
Db 259 SARQDFQREAEILTMLOHQHIVRFVCTEGRPLLMVFYEMRHGDLNRELSRSHGPDALL 318
QY 724 EGAPGCGQAAGGPTISYPMLLHVAQAISGMRYLATLNFVHRDLATRNCLVGENETIKIA 783
Db 319 AGGE---DVAPGP-LGLGQLLAVASQAAGVYLAGLHFVHRDLATRNCLVGGGLVVKIG 374
QY 784 DFGMSRLNAGDYRVOGRAVLPPIRMWAEICILMGKFTTASDVAFGVTLMWELMCLRAQ 843
Db 375 DFGMSRDIYDYRGGRTMLPIRMWPPESILYRKFTTESDVWSEFGVYLWEIFTYVK-Q 433
QY 844 PFQGLTDEQVIENAGFEFFRDQGVYLSRPPACPGQGLYELMLRCWSEQRPPFSQLH 902
Db 434 PWQLSNTAIDCI-----TQGRE--LERPRACPPVEYAIMRCWCQWREPPQORHSIKDVH 485

RESULT 12
Q15656 PRELIMINARY; PRT; 591 AA.
AC Q15656;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE P68 TRK-T3 ONCOPROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA GRECO A., MARIANI C., MIRANDA C., LUPAS A., PAGLIARDINI S., POMATI M.,
RX MEDLINE; 96025992.
RT novel gene on chromosome 3 whose product has a potential coiled-coil
RT domain.
RL Mol. Cell. Biol. 15:6118-6127(1995).
DR EMBL; X65960; CAA59936.1;
DR PFAM; PF00069; pkinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
SQ SEQUENCE 591 AA; 65851 MW; 8DAAFAC4 CRC32;

```

```

Query Match 12.8%; Score 627; DB 4; Length 591;
Best Local Similarity 45.5%; Pred. No. 1.le-44;
Matches 136; Conservative 49; Mismatches 84; Indels 30; Gaps 8;

```

```

QY 607 RSLRKEKLGEGQFVHLCEVDS---PQDLVSLDFPLNVRKGHPLLVAVKILRPDATK 663
Db 302 RDIVLKWELGEGAFGKVFLEASHNLLPEQD-----KMLVAVKALK-EASE 346
QY 664 NARNDLKEVKIMSLKDPNIIRLLGVCVQDDPLCMITDYMNGDNLQFSLAHLQEDKAA 723
Db 347 SARQDFQREAEILTMLOHQHIVRFVCTEGRPLLMVFYEMRHGDLNRELSRSHGPDALL 406
QY 724 EGAPGCGQAAGGPTISYPMLLHVAQAISGMRYLATLNFVHRDLATRNCLVGENETIKIA 783
Db 407 AGGE---DVAPGP-LGLGQLLAVASQAAGVYLAGLHFVHRDLATRNCLVGGGLVVKIG 462
QY 784 DFGMSRLNAGDYRVOGRAVLPPIRMWAEICILMGKFTTASDVAFGVTLMWELMCLRAQ 843
Db 463 DFGMSRDIYDYRGGRTMLPIRMWPPESILYRKFTTESDVWSEFGVYLWEIFTYVK-Q 521
QY 844 PFQGLTDEQVIENAGFEFFRDQGVYLSRPPACPGQGLYELMLRCWSEQRPPFSQLH 902
Db 522 PWQLSNTAIDCI-----TQGRE--LERPRACPPVEYAIMRCWCQWREPPQORHSIKDVH 573

RESULT 13
O43621 PRELIMINARY; PRT; 422 AA.
ID O43621

```

```

AC O43621;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DE ETS RELATED PROTEIN-GROWTH FACTOR RECEPTOR TYROSINE KINASE FUSION
DE PROTEINS (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominidae; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RA KNEZEVICH S.R., MCFADDEN D.E., TAO W., LIM J.F., SORESENSEN P.H.B.;
RL Nat. Genet. 18:0-0(1998).
DR EMBL; AF041811; AAC12728.1;
DR PFAM; PF00069; pkinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.
FT NON_TER 1
SQ SEQUENCE 422 AA; 48156 MW; D6165427 CRC32;

Query Match 12.7%; Score 625; DB 4; Length 422;
Best Local Similarity 42.6%; Pred. No. 1e-44;
Matches 143; Conservative 51; Mismatches 90; Indels 52; Gaps 11;

QY 590 ALPPGAVGDDGPPRVDFPRSLRKEKLGEGQFVHLCEVDSQDLVSLDFPLNVRKGHP 649
Db 104 AMPIGRTADVQ---HIKRRDVLARELGEGAFGKVFLEAC-----YNLSPTK-DK 149
QY 650 LLVAVKILRPDATKNARNDLKEVKIMSLKDPNIIRLLGVCVQDDPLCMITDYMNGDNL 709
Db 150 MLVAVKALK-DPTLAARKDKQREAEILTMLOHQHIVRFVCTEGRPLLMVFYEMRHGDL 208
QY 710 NQFLSAHQLEDKAEAGAPG-----DGOAAGP-TISTPMLLHVAQAISGMRYLATLNFY 763
Db 209 NKFLRAH-----GPDAMILVDGPPROAKGELGSLQMLHIASQIASGMVYLASOHFV 259
QY 764 HRDLATRNCLVGENFTIKIADFGMSRLNLYAGDYR-----VQGRAVLPIRW 809
Db 260 HRDLATRNCLVGENFTIKIADFGMSRLNLYAGDYR-----VQGRAVLPIRW 319
QY 810 MAWECILMGKFTTASDVAFGVTLMWELMCLRAQFPQGLTDEQVIENAGFEFFRDQGVY 869
Db 320 MPPEISIMYRKFTTESDVWSEFGVYLWEIFTYVK-QPWFQLSNTVEIECI-----TQGR--V 371
QY 870 LSRPPACPGQGLYELMLRCWSEQRPPFSQLHRL 905
Db 372 LERPRVCPKEYDYVMLGCGWQREPPQORLNKEIYKIL 407

RESULT 14
Q91373 PRELIMINARY; PRT; 282 AA.
ID Q91373
AC Q91373;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE TRKB (FRAGMENT).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae; Xenopodinae;
RN [1]
RP SEQUENCE FROM N.A.
RA COHEN-CORY S., FRASER S.E.;
RT "BDNF in the development of the visual system of Xenopus.";
RL Neuron 12:747-761(1994).
DR EMBL; S69713; AAB30791.1;
DR PFAM; PF00069; pkinase; 1.
DR PROSITE; PS00239; RECEPTOR_TYR_KIN_II; 1.

```

FT NON\_TER 1 1  
SQ SEQUENCE 282 AA; 32820 MW; AA804CBA CRC32;

Query Match 12.78; Score 621; DB 13; Length 282;  
Best Local Similarity 45.78; Pred. No. 1.3e-44;  
Matches 132; Conservative 49; Mismatches 74; Indels 34; Gaps 8;

QY 607 RSRLRFKEKLGEGQGEVHLCVEVDSQDLVSLDFPLNV-RKGHPHLLVAVKILRPDANKNA 665  
DB 24 RHINIVLKRELGEGAFKGVFLAEC-----YNLRYEQDKILVAVKTLK-DASDNA 70  
QY 666 RNDLFKEVKIMSRKLPDNIIRLLGVCVODDPLCMITDYMENGDLNQFLSAHQLEDKAAEG 725  
DB 71 RKDFHREAELLTNLQHENIVKFGVCGVDPLIMVFYMKHGDNLKELRAH-----122  
QY 726 APGDCQAAQGP---TISYPMLLHVAAQIATSGMRYLATLNFVHRDLATNCLVGENFTIKI 782  
DB 122 GPDAVMAEGNLLIELTQSMIHISQQAAGMXYLASQHFVHRDLPTNCLVGENLLVKI 181  
QY 783 ADFGMSRLYAGDYRVQGRAVLPPIRMMAWECILMGKFTTASDVYAFGVTLMVEVLMCLRA 842  
DB 182 GDFGMSRDVSYDYRVGHTMLPIRMWPPESIMYRRTTESDVWSLGVLMVEIFTYK- 241  
QY 843 QPFGLTDEQVIENAGEFFRDQGRQVYLSRPPACPGQGLYELMLRCWSRE 891  
DB 241 QPWYQLSNNEVIECI-----TQGR--VLQRPRTCPKEIYDLMRGQWORE 282

RESULT 15  
Q17576 PRELIMINARY; PRT; 806 AA.  
AC Q17576;  
DT 01-NOV-1996 (TRENBLrel. 01, Created)  
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)  
DE KIN-8 PROTEIN.  
GN KIN-8.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;  
OC Rhabditina; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA BERKS M.;  
RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE; 94150718.  
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,  
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,  
RA CRAYTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,  
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,  
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,  
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,  
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,  
RA SMAILDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,  
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,  
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RT elegans.";  
RL Nature 368:32-38(1994).  
DR EMBL; Z35595; CAA84639.1; -;  
DR PFAM; PF00051; kring1; 1.  
DR PFAM; PF00069; pkinase; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
SQ SEQUENCE 806 AA; 90384 MW; 8505D993 CRC32;

Query Match 12.68; Score 618; DB 5; Length 806;  
Best Local Similarity 36.28; Pred. No. 1e-43;  
Matches 151; Conservative 55; Mismatches 151; Indels 60; Gaps 7;

QY 485 QBRPRGNPPHSPACVPNGSALLSNPAYRLLLATYARPPRGPPPTPAWAKPTNTQAYS 544  
DB 375 QKTRHONAHCSAPSINSAA-----NSAYYRKLNGTSTPIMGRVPPHVE-----420  
QY 545 GYMEPEKFCAPLLPPPPQNSVPHYAEADIVTLQGVTCGNTYAVPALPGAVGDGPPRVD 604  
DB 420 ---MTSLPSAHLGPPPPYPMQHLQOARRFPSPQPIDDINSYKVEFITP-----466  
QY 605 FPRSRLRFKEKLGEGQGEVHLCVEVDSQDLVSLDFPLNVKRGHPHLLVAVKILRPDANKNA 664  
DB 466 ---SOLSVREKIGEGQGVVH-----SGIYTSGLFAPEPMVAVKKCRHDATNA 511  
QY 665 ARNDFLKEVKIMSRKLPDNIIRLLGVCVODDPLCMITDYMENGDLNQFLSAHQLEDKAAE 724  
DB 512 ERAQLQEITRAVATEDHPNVIKLIGVCYMDNSLLAVFEYVMVHGDLHELLKVR-----564  
QY 725 GAPGDCQAAQGP---TISYPMLLHVAAQIATSGMRYLATLNFVHRDLATNCLVGENFTIKI 784  
DB 564 -VPPADHDMGGITEANAEFLYIATQALGMEYLASMSFVHRDLATNCLVGDTRIKIAD 622  
QY 785 FGMRSRLYAGDYRVQGRAVLPPIRMMAWECILMGKFTTASDVYAFGVTLMVEVLMCLRAQ 844  
DB 623 FGLMRTSYGSDYYKMLHRSWMPVRWMSKEAIEQGRFSEASDVWSFGVTLMVEIWSFGR-QP 681  
QY 845 FGQLTDEQVIENAGEFFRDQGRQVYLSRPPACPGQGLYELMLRCWSRESEQRPPFSQL 901  
DB 682 YEGASNQVIELV-----ANRHLLECPHNCPTNIYSLMVECMHENIERRPTFSEI 731

Search completed: November 4, 1999, 01:53:54  
Job time: 7614 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 1999, 23:45:30 ; Search time 18.36 Seconds  
(without alignments)  
568.318 Million cell updates/sec

Title: US-08-170-558-4  
Perfect score: 4903  
Sequence: 1 MGPEALSSLLLLLLVASGDA.....QRPFFSOLHRFLAEDALNTV 913

Scoring table: BLOSUM62

Searched: 119832 seqs, 11428610 residues

Database : Issued Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/PCTUS9\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4903	100.0	913	1	US-08-445-640-4
2	4882	99.6	919	1	US-08-336-343A-2
3	2415	49.3	855	1	US-08-336-343A-4
4	2413	49.2	854	2	US-08-456-647B-20
5	2413	49.2	854	2	US-08-237-401A-20
6	2175	44.4	399	1	US-08-445-640-8
7	1680	34.3	317	2	US-08-701-191A-25
8	659.5	13.5	821	1	US-08-339-578-2
9	659	13.4	825	2	US-08-469-537A-73
10	658	13.4	847	2	US-08-286-305A-5
11	658	13.4	822	2	US-08-359-705B-2
12	658	13.4	822	2	US-08-286-846A-2
13	658	13.4	847	2	US-08-441-104A-5
14	658	13.4	822	2	US-08-457-880A-2
15	658	13.4	847	2	US-08-440-816A-5
16	656	13.4	279	2	US-08-469-537A-51
17	650.5	13.3	850	2	US-08-286-305A-7
18	650.5	13.3	850	2	US-08-441-104A-7
19	650.5	13.3	850	2	US-08-440-816A-7
20	638	13.0	814	2	US-08-286-305A-3
21	638	13.0	790	2	US-08-359-705B-9
22	638	13.0	790	2	US-08-286-846A-9
23	638	13.0	814	2	US-08-441-104A-3
24	638	13.0	790	2	US-08-457-880A-9
25	638	13.0	814	2	US-08-440-816A-3
26	633.5	12.9	839	2	US-08-359-705B-6
27	633.5	12.9	839	2	US-08-286-846A-6
28	633.5	12.9	839	2	US-08-457-880A-6
29	632.5	12.9	825	1	US-07-912-952-2
30	630	12.8	281	2	US-08-469-537A-52
31	628	12.8	310	2	US-08-469-537A-74
32	628	12.8	294	2	US-08-701-191A-20
33	627	12.8	728	1	US-07-912-952-4
34	620.5	12.7	268	2	US-07-857-224B-74
35	608.5	12.4	946	3	PCT-US95-08493-13
36	606.5	12.4	868	1	US-08-374-834-1
37	606.5	12.4	868	2	US-08-644-271-1
38	603	12.3	868	3	PCT-US95-08493-21
39	601.5	12.3	869	1	US-08-374-834-16

40 601.5 12.3 869 2 US-08-644-271-29  
41 -600.5 12.2 530 3 PCT-US95-08493-2  
42 600.5 12.2 860 3 PCT-US95-08493-19  
43 599 12.2 304 2 US-08-701-191A-30  
44 546 11.1 800 2 US-08-469-537A-72  
45 546 11.1 800 2 US-08-469-537A-78

## ALIGNMENTS

RESULT 1  
US-08-445-640-4  
; Sequence 4, Application US/08445640  
; Patent No. 5709858  
; GENERAL INFORMATION:  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Mark, Melanie R.  
; APPLICANT: Scadden, David T.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Baron, Will F.  
; TITLE OF INVENTION: Protein Tyrosine Kinases  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/445,640  
; FILING DATE: 22-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/170558  
; FILING DATE: 20-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/157563  
; FILING DATE: 23-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hasak, Janet E.  
; REGISTRATION NUMBER: 28,616  
; REFERENCE/DOCKET NUMBER: 854C2  
; TELEPHONE: 415/225-1896  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 913 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
US-08-445-640-4

Query Match 100.0%; Score 4903; DB 1; Length 913;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 913; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGPEALSSLLLLLLVASGDADMKGHFDPKACRYALGMDQRTIPDSDISASSWSNSTAAR 60  
Db 1 MGPEALSSLLLLLLVASGDADMKGHFDPKACRYALGMDQRTIPDSDISASSWSNSTAAR 60  
QY 61 HSRLESSDGDGAWCAGSVFPKEEYLVQDQLRHLVALVGTQGRHAGLGKFSRSL 120  
Db 61 HSRLESSDGDGAWCAGSVFPKEEYLVQDQLRHLVALVGTQGRHAGLGKFSRSL 120

```

QY 121 RYSDGRWGWKDRGQEVISGNEDEGVVVKDLGPPMVARLVRYPRADRVMSVCLRV 180
Db 121 RYSDGRWGWKDRGQEVISGNEDEGVVVKDLGPPMVARLVRYPRADRVMSVCLRV 180
QY 181 ELYGCLWRDGLLSYAPVGTMYLSEAVYLNDSYDGHVGVGLQGLADGVVGLDD 240
Db 181 ELYGCLWRDGLLSYAPVGTMYLSEAVYLNDSYDGHVGVGLQGLADGVVGLDD 240
QY 241 FRKSOELRVWPGDYVYVWNSHSSGYSYVEMEFDFLRFAQMVHNCNMHTLGARLPGG 300
Db 241 FRKSOELRVWPGDYVYVWNSHSSGYSYVEMEFDFLRFAQMVHNCNMHTLGARLPGG 300
QY 301 VECFRFRGPAMAWEGEPMRHNLGNLGDPRARAVSVPLGGRVARFLQCRFLFAGPWLIFS 360
Db 301 VECFRFRGPAMAWEGEPMRHNLGNLGDPRARAVSVPLGGRVARFLQCRFLFAGPWLIFS 360
QY 361 EISFISDVVNSSPALGGTFPPAPWPPGPPPTNFSSLELEPRGQOPVAKPEGSTAILI 420
Db 361 EISFISDVVNSSPALGGTFPPAPWPPGPPPTNFSSLELEPRGQOPVAKPEGSTAILI 420
QY 421 GCLVAIILLLLIIALMLRHLWRLLSKAERRVLEELTVHLSVPGDTILINNRPGPRE 480
Db 421 GCLVAIILLLLIIALMLRHLWRLLSKAERRVLEELTVHLSVPGDTILINNRPGPRE 480
QY 481 PPPYQEPFRGNPPHSPACVPNGSALLSNPAYRLLLATYARPPRGPPPTPAWAKPTNT 540
Db 481 PPPYQEPFRGNPPHSPACVPNGSALLSNPAYRLLLATYARPPRGPPPTPAWAKPTNT 540
QY 541 QAYSGDYMEPEKPGAPLLPPPPONSYPHYAEADIVTLQGVGTGNTYAVPALPGGAVGDP 600
Db 541 QAYSGDYMEPEKPGAPLLPPPPONSYPHYAEADIVTLQGVGTGNTYAVPALPGGAVGDP 600
QY 601 PRVDFPRSRRLRFEKLGEGQFGEVHLCVDSQDLYSLDFFPLNVRKGHLLVAVKILRPD 660
Db 601 PRVDFPRSRRLRFEKLGEGQFGEVHLCVDSQDLYSLDFFPLNVRKGHLLVAVKILRPD 660
QY 661 ATKARNRDFLKEVKIMSRKDPNIIRLLGVCVODDPLCMITDMYNGDNLQFLSAHLQED 720
Db 661 ATKARNRDFLKEVKIMSRKDPNIIRLLGVCVODDPLCMITDMYNGDNLQFLSAHLQED 720
QY 721 KAEGAPGDCQAAGPTTISYPMLLHVAAQIATSGMRYLATLNFVHRDLATRNCLVGENFTI 780
Db 721 KAEGAPGDCQAAGPTTISYPMLLHVAAQIATSGMRYLATLNFVHRDLATRNCLVGENFTI 780
QY 781 KIADEGSRNLYAGDYRVQGRAVLPIRMAWECILMGKFTTASDYWAFGVTILWEVIMLC 840
Db 781 KIADEGSRNLYAGDYRVQGRAVLPIRMAWECILMGKFTTASDYWAFGVTILWEVIMLC 840
QY 841 RAQPFGLTDEQVIENAGEFRDQGRQVYLSRPPACQGLYELMLRCWSRESEQRPPFSQ 900
Db 841 RAQPFGLTDEQVIENAGEFRDQGRQVYLSRPPACQGLYELMLRCWSRESEQRPPFSQ 900
QY 901 LHRELAEDALNTV 913
Db 901 LHRELAEDALNTV 913

```

```

RESULT 2
US-08-336-343A-2
: Sequence 2, Application US/08336343A
: Patent No. 5677144
: GENERAL INFORMATION:
: APPLICANT: Ullrich, Axel
: APPLICANT: Alves, Frauke
: TITLE OF INVENTION: CKK-2, A No. 5677144el Receptor Tyrosine Kinase
: NUMBER OF SEQUENCES: 43
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10036-2711

```

```

: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/336,343A
: FILING DATE: 08-NOV-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Coruzzi, Laura A.
: REGISTRATION NUMBER: 30,742
: REFERENCE/DOCKET NUMBER: 7683-065
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090
: TELEFAX: (212) 869-9741/8864
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 919 amino acids
: TYPE: amino acid
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: US-08-336-343A-2

```

```

Query Match 99.6%; Score 4882; DB 1; Length 919;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 912; Conservative 0; Mismatches 1; Indels 6; Gaps 1;

QY 1 MGPEALSSLLLLLVASGDADMGKGFDPKCRYALGMDRTIPDSISASSSSDSTAAR 60
Db 1 MGPEALSSLLLLLVASGDADMGKGFDPKCRYALGMDRTIPDSISASSSSDSTAAR 60
QY 61 HSRLESSDGDGAWCPAGSVPEKEEYLQVDLQRLHLVALVGTGRHAGGLGKEFSRYRL 120
Db 61 HSRLESSDGDGAWCPAGSVPEKEEYLQVDLQRLHLVALVGTGRHAGGLGKEFSRYRL 120
QY 121 RYSDGRWGWKDRGQEVISGNEDEGVVVKDLGPPMVARLVRYPRADRVMSVCLRV 180
Db 121 RYSDGRWGWKDRGQEVISGNEDEGVVVKDLGPPMVARLVRYPRADRVMSVCLRV 180
QY 181 ELYGCLWRDGLLSYAPVGTMYLSEAVYLNDSYDGHVGVGLQGLADGVVGLDD 240
Db 181 ELYGCLWRDGLLSYAPVGTMYLSEAVYLNDSYDGHVGVGLQGLADGVVGLDD 240
QY 241 FRKSOELRVWPGDYVYVWNSHSSGYSYVEMEFDFLRFAQMVHNCNMHTLGARLPGG 300
Db 241 FRKSOELRVWPGDYVYVWNSHSSGYSYVEMEFDFLRFAQMVHNCNMHTLGARLPGG 300
QY 301 VECFRFRGPAMAWEGEPMRHNLGNLGDPRARAVSVPLGGRVARFLQCRFLFAGPWLIFS 360
Db 301 VECFRFRGPAMAWEGEPMRHNLGNLGDPRARAVSVPLGGRVARFLQCRFLFAGPWLIFS 360
QY 361 EISFISDVVNSSPALGGTFPPAPWPPGPPPTNFSSLELEPRGQOPVAKPEGSTAILI 420
Db 361 EISFISDVVNSSPALGGTFPPAPWPPGPPPTNFSSLELEPRGQOPVAKPEGSTAILI 420
QY 421 GCLVAIILLLLIIALMLRHLWRLLSKAERRVLEELTVHLSVPGDTILINNRPGPRE 480
Db 421 GCLVAIILLLLIIALMLRHLWRLLSKAERRVLEELTVHLSVPGDTILINNRPGPRE 480
QY 481 PPPYQEPFRGNPPHSPACVPNGSALLSNPAYRLLLATYARPPRGPPPTPAWAKPTNT 540
Db 481 PPPYQEPFRGNPPHSPACVPNGSALLSNPAYRLLLATYARPPRGPPPTPAWAKPTNT 540
QY 541 QAYSGDYMEPEKPGAPLLPPPPONSYPHYAEADIVTLQGVGTGNTYAVPALPGGAVGDP 600
Db 541 QAYSGDYMEPEKPGAPLLPPPPONSYPHYAEADIVTLQGVGTGNTYAVPALPGGAVGDP 600
QY 601 PRVDFPRSRRLRFEKLGEGQFGEVHLCVDSQDLYSLDFFPLNVRKGHLLVAVKILRPD 660
Db 601 PRVDFPRSRRLRFEKLGEGQFGEVHLCVDSQDLYSLDFFPLNVRKGHLLVAVKILRPD 660

```



COUNTRY: US  
 ZIP: 92037  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/456,647B  
 FILING DATE: 02-JUN-1995  
 CLASSIFICATION: 530  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/237,401  
 FILING DATE: 02-MAY-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/884,486  
 FILING DATE: 15-MAY-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Wetherell Ph.D., John R.  
 REGISTRATION NUMBER: 31,678  
 REFERENCE/DOCKET NUMBER: 07251/007002  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (619) 678-5070  
 TELEFAX: (619) 678-5099  
 INFORMATION FOR SEQ ID NO: 20:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 854 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-456-647B-20

Query Match 49.2%; Score 2413; DB 2; Length 854;  
 Best Local Similarity 52.3%; Pred. No. 1e-166;  
 Matches 481; Conservative 119; Mismatches 220; Indels 100; Gaps 15;

QY 9 LLLLLLVSAGDADMGHFDPAKRYALGMQDRTIPDSISASSWSDSTAARHSLRSSD 68  
 DB 10 VLLLLLLIUGSA--KAQNPATCRPLGMSGHIPDEIDTASSQWSESTAARYKRLDSEE 67

QY 69 GDGACPCAGSVFPKPE-EYLVQDQLRLHLVALVGTQGRHAGLGLKEFSRSLRYSDGR 127  
 DB 68 GDGACPCETVPQDDLKLEFLQIDRLTLFTLVTGQGRHAGHGFIEFAPMYKINYSRDS 127

QY 128 RMWKKDRWGQGVISGNDPEGVKLDLGPMPVARLVRYFPRADRVMSVCLRVLYGLCW 187  
 DB 128 RWISNRHKGKOVLDGNSNPYDFLKLDPPIVAFVRLIPVTHSMNVCMRVLYGCWV 187

QY 188 RDGLLSYAPVGTMYL--SEAVYINDSTYDGHVGLQGLQGLADGVVGLDDFRKSQ 245  
 DB 188 LDGLVSNAPAGQQFVLPGGSIYIYNDSDYDG-AVGSYMTGLGLTQDGVSLDQDTQTH 246

QY 246 ELRVNPGDYVGVWSSHSRSGVVEFEFDRLRARQAMQVHCNNMHTLGARLPGVCECF 305  
 DB 247 EYHVNPGDYVGVWSSHSRSGVVEFEFDRLRARQAMQVHCNNMHTLGARLPGVCECF 306

QY 306 RGPAMAWGEPNRLNGLNGLDPRARAVSVPLGGRVARFQCRFLFAGPWLLESEISFI 365  
 DB 307 -RSEASENEPTAVYPLVLDVNPASRVTVPLHHRMASAKCOVHFADTWMPSEITFQ 365

QY 366 SD--VYVNSPALGGTFFPAPWPPGPPPTNFSSLELPRGQPVAKPEGSPTAILIGCL 423  
 DB 366 SDAAMYNN--GALPTSP-----MAPTYDPMKLVDDSDNTRILIGCL 405

QY 424 VAILLLLLLIALMLWHLHRLLSKAERVLLEELTVHLSVPGDTILINNR---PGPR 479  
 DB 406 VAIIFILLAIIVILWQVQWLEKASRRMLDDEMTVSLSPSSSEMFNNRRSSPSEQ 465

QY 480 EP-----PPYQEPNRPNGPHSPACPVNGSALILSNPAYRLLLATYARPPRPG 528  
 DB 466 ESNSTYDRIFPLRPDQEP-----SRLIRKLPEF----- 495

QY 529 PPTPAWAKPTNTQAYSGDYMEPEKPGAPLLPPPPQNSVPHYAEADIVTLOGVTGNTYAV 588  
 DB 495 -----APGEEESCSCGVVKAQPNGP-----EGVPHYAEADIVNLOGVTGNTYCV 540

QY 589 PALPGAVGDGPPRV-DPPRSRLRFKEKELGEGQFGEVHLCFVDSQDVLVSLDFPLNVRKG 647  
 DB 541 PAVTMDLISGKDVAVEEPKRLAFKEKELGEGQFGEVHLCFVDSQDVLVSLDFPLNVRKG 600

QY 648 HPLLVAVKILRPDANKNARNDLFKEVKIMSKLPNIIRLLGVGVQDDPLCMITDYMWNG 707  
 DB 601 QPVLVAVKMLRADANKNARNDLFKEIKIMSKLPNIIRLLAVCITEDPLCMITEYMWNG 660

QY 708 DLNQFLSAHLEDKAAEGAPGDQQAQPTTISYPMLLHVAQAQIASGMYRLATLNFVHRDL 767  
 DB 661 DLNQFLSRHREPLSSCSDA-----TVSYANLKFMTQIASGMYRLATLNFVHRDL 710

QY 768 ATRNCLVGENFTIKIADFGMSRLNLYAGDYRVQGRAVLPIRMMAWECILMGKFTTASDVW 827  
 DB 711 ATRNCLVGNKTIKADFGMSRLNLYAGDYRVQGRAVLPIRMMAWECILMGKFTTASDVW 770

QY 828 AFGVTLMWVLMCLRAQFQGLTDEQVIEVNAEFGFRDQGRQVYLSRPPACPOGLVLMRLC 887  
 DB 771 AFGVTLMWVLMCLRAQFQGLTDEQVIEVNAEFGFRDQGRQVYLSRPPACPOGLVLMRLC 830

QY 888 WSRSEQRPPFSQLHRFLAE 907  
 DB 831 WRRETKHRPESQEIHLHLQ 850

RESULT 5  
 US-08-237-401A-20  
 Sequence 20, Application US/08237401A  
 Patent No. 5837448  
 GENERAL INFORMATION:  
 APPLICANT: Lemke Ph.D. et al., Greg E.  
 TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES  
 NUMBER OF SEQUENCES: 54  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Fish & Richardson P.C.  
 STREET: 4225 Executive Square, Suite 1400  
 CITY: La Jolla  
 STATE: CA  
 COUNTRY: US  
 ZIP: 92037  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/237,401A  
 FILING DATE: 02-MAY-1994  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/884,486  
 FILING DATE: 15-MAY-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Haile Ph.D., Lisa A.  
 REGISTRATION NUMBER: 38,347  
 REFERENCE/DOCKET NUMBER: 07251/007001  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (619) 678-5070  
 TELEFAX: (619) 678-5099  
 INFORMATION FOR SEQ ID NO: 20:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 854 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-237-401A-20

Query Match 49.2%; Score 2413; DB 2; Length 854;





QY 379 TFPAPWPPGPPPTNFSSLEPRGQOPVAKPEGSPTA 417  
 Db 361 TFPAPWPPGPPPTNFSSLEPRGQOPVAKPEGSPTA 399

RESULT 7

US-08-701-191A-25  
 ; Sequence 25, Application US/08701191A  
 ; Patent No. 5942428  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Moosa Mohammadi, Joseph Schlessinger,  
 ; APPLICANT: and Stevan R. Hubbard  
 ; TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN  
 ; TITLE OF INVENTION: OF NON-INSULIN RECEPTOR TYROSINE KINASE  
 ; NUMBER OF SEQUENCES: 41  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Lyon & Lyon  
 ; STREET: 633 West Fifth Street  
 ; STREET: Suite 4700  
 ; CITY: Los Angeles  
 ; STATE: California  
 ; COUNTRY: U.S.A.  
 ; ZIP: 90071-2066  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 ; MEDIUM TYPE: storage  
 ; COMPUTER: IBM compatible  
 ; OPERATING SYSTEM: IBM P.C. DOS 5.0  
 ; SOFTWARE: FASTSEQ for Windows 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08701.191A  
 ; FILING DATE: August 21, 1996  
 ; CLASSIFICATION: 530  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER:  
 ; FILING DATE:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Warburg, Richard J.  
 ; REGISTRATION NUMBER: 32,327  
 ; REFERENCE/DOCKET NUMBER: 227/088  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (213) 489-1600  
 ; TELEFAX: (213) 955-0440  
 ; TELEX: 67-3510  
 ; INFORMATION FOR SEQ ID NO: 25:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 317 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-701-191A-25

Query Match 34.3%; Score 1680; DB 2; Length 317;  
 Best Local Similarity 99.7%; Pred. No. 2.8e-114;  
 Matches 316; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 597 GGGPPRVDFPRSLRFEKLGEGQFGEVHLCEVDSPQDLVSLDFPLNVRKGGHPLLVAVKI 656  
 Db 1 GGGPPRVDFPRSLRFEKLGEGQFGEVHLCEVDSPQDLVSLDFPLNVRKGGHPLLVAVKI 60  
 QY 657 LRPDATKNARNDFLEKVKIMSKLPNIIIRLLGVCVQDDPLCMITDYMGDNLQFLSAH 716  
 Db 61 LRPDATKNARNDFLEKVKIMSKLPNIIIRLLGVCVQDDPLCMITDYMGDNLQFLSAH 120  
 QY 717 QLEDKAAEGAPGQQAQGTTSYPMMLHVAQAISGMRYLATLNFVRDLATRNCLVGE 776  
 Db 121 QLEDKAAEGAPGQQAQGTTSYPMMLHVAQAISGMRYLATLNFVRDLATRNCLVGE 180  
 QY 777 NFTIKIADFGMSRLNLYAGDYRVQGRAVLPINRWAECEILMGKFTTASDVWAFGVTLWEV 836  
 Db 181 NFTIKIADFGMSRLNLYAGDYRVQGRAVLPINRWAECEILMGKFTTASDVWAFGVTLWEV 240

QY 837 LMLCRAQPGQLTDEQVIENAGEFFRDQGRVYLSRPPACPGQLYELMLRCWSESEQRP 896  
 Db 241 LMLCRAQPGQLTDEQVIENAGEFFRDQGRVYLSRPPACPGQLYELMLRCWSESEQRP 300  
 QY 897 PFSQLHRLFLAEDALNTV 913  
 Db 301 PFSQLHRLFLAEDALNTV 317

RESULT 8

US-08-339-578-2  
 ; Sequence 2, Application US/08339578  
 ; Patent No. 5622862  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Squinto, et al.  
 ; TITLE OF INVENTION: ASSAY SYSTEMS FOR NEUROTROPHIN ACTIVITY  
 ; NUMBER OF SEQUENCES: 2  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Regeneron Pharmaceuticals, Inc.  
 ; STREET: 777 Old Saw Mill River Road  
 ; CITY: Tarrytown  
 ; STATE: New York  
 ; COUNTRY: U.S.A.  
 ; ZIP: 10591-6707  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/339,578  
 ; FILING DATE: 14-NOV-1994  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/690,199  
 ; FILING DATE: 23-APR-1991  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Kempler, Gail M.  
 ; REGISTRATION NUMBER: 32,143  
 ; REFERENCE/DOCKET NUMBER: 6526-061A  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (914) 345-7400  
 ; TELEFAX: (914) 345-7721  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 821 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-339-578-2

Query Match 13.5%; Score 659.5; DB 1; Length 821;  
 Best Local Similarity 30.4%; Pred. No. 5.2e-40;  
 Matches 201; Conservative 72; Mismatches 186; Indels 201; Gaps 22;  
 QY 373 SPALGTFPPAPWPPGPPPTNFSSLEPRGQOPVAKPEGSPTAILICGLVAILLLLL 432  
 Db 217 SCVSGDPLTLYWDYGNLVSKHNETSHTQGLRITNISSDDSGKQISCV----- 268  
 QY 433 IIALMLWLHWRLLSKAERRVLEELTVHLSV---PGDTILINNRPGRPPEP----- 483  
 Db 268 -----AENLVGEDQDSVNLTVHFAPTITFL-----ESPTSDHHC 302  
 QY 483 -PYOEPRPGNPHSPCPVNGS-----ALLLSNPA-----Y 513  
 Db 303 IPP---TVRGNPDLQWFNGYNGAILNESKYICTKIHTVNHTEHGLQLDNPHMNGDY 359  
 QY 514 RLL-----ATYARPPRGPPPTPAWAKPTNTQAYSGDYMEP--E 551  
 Db 360 TLMKNEYGKDERQISAHEFMGRPGVDYETPNPNPEVLYEDWTPTDTI----GDTTKNSNE 415

QY 552 KPGAPLLPPPPQNSVPHYAADIIVTLQV----- 581  
Db 416 IPSTVDADQSNREHLSVYAVVIVASVVGFCLLVLMLLLLKLAHRSFGMKGPASVINSDD 475  
QY 581 -----TGGNTYAVPALPGAVGDGP---PRVDFP-----RS 608  
Db 476 SASPLHHISNGSNTPSSSEGGPDVILGKIPVIENPOYFGITNSQLKPDTFVQHKKR 535  
QY 609 RLRFKELGEGGFGVHLC-----VDSPODLSLDFPLNVRKGHPLLVAVKILRPDATKNA 665  
Db 536 NIIVKRELGEAGFKVFLAECYNLCPEQD-----KILVAVKTLK-DASDNA 580  
QY 666 RNDLFKEVIMSLKDPNIIRLLGVCVODDPLCMITDYMNGDNLNQLSAHQLED-KAAE 724  
Db 581 RKFHREALLNLQHEHIVKFGVGVGDPILIMVFEYMKHGLKFLRAHGPDAVLMAE 640  
QY 725 GAPGDQAAQAGPT-ISTPMLLHVAQAIAAGMRYLATLNFVHRDLATNCLVGENFTIKIA 783  
Db 641 GNP-----PTELTQSOMLHIAQIAAGVYLASQHFVHRDLATNCLVGENLLVKIG 692  
QY 784 DFGMSRLNLYAGDYRYVQGRVPLPIRMWAWECILMGKFTTASDVWAFVTLWEVLMCLRAQ 843  
Db 693 DFGMSRDVYSTDYRYVGGHTMLPIRMWPPESIMYRKFTTESDVWLSGLVWEIFYTK-Q 751  
QY 844 PFGQLTDEQVIENAGEFFRDGROVYLSRPPACQGLYELMLRCWSRESORPPFSOLHR 903  
Db 752 PWYQLSNNEVIECI-----TQGR--VLQRPRTCPQEVYELMLGCMQREPHTRKNKSIHT 804  
QY 904 FL 905  
Db 805 LL 806

## RESULT 9

US-08-469-537A-73  
; Sequence 73, Application US/08469537A  
; Patent No. 5843749  
; GENERAL INFORMATION:  
; APPLICANT: Maisonnier, et al.  
; TITLE OF INVENTION: EHK AND ROR TYROSINE  
; TITLE OF INVENTION: KINASES  
; NUMBER OF SEQUENCES: 107  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Regeneron Pharmaceuticals, Inc.  
; STREET: 777 Old Saw Mill River Road  
; CITY: Tarrytown  
; STATE: NY  
; COUNTRY: U.S.A.  
; ZIP: 10591  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/469,537A  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: USN 08/406,247  
; FILING DATE: 17-MAR-1995  
; APPLICATION NUMBER: USN 08/144,992  
; FILING DATE: 28-OCT-1993  
; APPLICATION NUMBER: USN 07/736,559  
; FILING DATE: 26-JUL-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kempler, Ph.D., Gail M  
; REGISTRATION NUMBER: 32,143  
; REFERENCE/DOCKET NUMBER: REG 070C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 914-345-7400  
; TELEFAX: 914-345-7721  
; TELEX:

; INFORMATION FOR SEQ ID NO: 73:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 285 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-469-537A-73

Query Match 13.4%; Score 659; DB 2; Length 285;  
Best Local Similarity 47.4%; Pred. No. 1.5e-40;  
Matches 144; Conservative 45; Mismatches 79; Indels 36; Gaps 9;

QY 607 RSLRFKELGEGGFGVHLC-----VDSPODLSLDFPLNVRKGHPLLVAVKILRPDATK 663  
Db 11 RHNIIVKRELGEAGFKVFLAECYNLCPEQD-----KILVAVKTLK-DASD 55  
QY 664 NARNDLFKEVIMSLKDPNIIRLLGVCVODDPLCMITDYMNGDNLNQLSAHQLED-KA 722  
Db 56 NARDFHREALLNLQHEHIVKFGVGVGDPILIMVFEYMKHGLKFLRAHGPDAVL 115  
QY 723 AEGAPGDQAAQAGPT-ISTPMLLHVAQAIAAGMRYLATLNFVHRDLATNCLVGENFTIK 781  
Db 116 AEGNP-----PTELTQSOMLHIAQIAAGVYLASQHFVHRDLATNCLVGENLLVK 167  
QY 782 IADFGMSRLNLYAGDYRYVQGRVPLPIRMWAWECILMGKFTTASDVWAFVTLWEVLMCL 841  
Db 168 IGDFGMSRDVYSTDYRYVGGHTMLPIRMWPPESIMYRKFTTESDVWLSGLVWEIFYTK 227  
QY 842 AQPFGQLTDEQVIENAGEFFRDGROVYLSRPPACQGLYELMLRCWSRESORPPFSOL 901  
Db 228 -QPWYQLSNNEVIECI-----TQGR--VLQRPRTCPQEVYELMLGCMQREPHTRKNKI 279  
QY 902 HRFL 905  
Db 280 HTLL 283

## RESULT 10

US-08-286-305A-5  
; Sequence 5, Application US/08286305A  
; Patent No. 5766863  
; GENERAL INFORMATION:  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Mark, Melanie R.  
; APPLICANT: Sadtick, Michael D.  
; APPLICANT: Shelton, David L.  
; APPLICANT: Wong, Wai Lee Tan  
; TITLE OF INVENTION: KINASE RECEPTOR ACTIVATION ASSAY  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/286,305A  
; FILING DATE: 05-AUG-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/170558  
; FILING DATE: 20-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/157563  
; FILING DATE: 23-NOV-1993

```
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 854C1P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 847 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-286-305A-5

Query Match 13.4%; Score 658; DB 2; Length 847;
Best Local Similarity 47.4%; Pred. No. 6.9e-40;
Matches 144; Conservative 45; Mismatches 79; Indels 36; Gaps 9;

QY 607 RSLRFEKELGEGQGEVHLCE---VDSPODLVSLDFPLNVRKGHPLLVAVKILRPDATK 663
DB 560 RHNIYVKRELGEAGFGKVFLEACYNLCPEQD-----KILVAVKTLK-DASD 604

QY 664 NARNDLFKEVKIMSLKDPNIIRLLGVCVODDPLCMITDYMENGLNOFLSAHQLED-KA 722
DB 605 NARKDFHREAELLTNLQHEHIVKFGVCVEGDPPLIMVFYMKHGLNKLFLRAHGPDAVLM 664

QY 723 AEGAPGDGQAAQGPT-ISYPMLLHVAQAISGMRYLATLNFVHRDLATRNCLVGENFTIK 781
DB 665 AEGNP-----PTELTSQMLHIAQIAAGMVLASQHFVHRDLATRNCLVGENLLVK 716

QY 782 IADFGMSRLNLYAGDYRYVQGRAVLPKRWMAWECILMGKFTTASDWAFGVTLWEVLMCLR 841
DB 717 IGDFGMSRDYSTDYRYVGGHTMLPIRWMPPEIMRYKFTTESDWLSGVVLEIFYGK 776

QY 842 AQPFGOLTDEQVIENAGEFFRDQGRQVYLSRPPACPGQLYELMLRCWSRESEQRPPFSOL 901
DB 777 -QPWYQLSNNVEIECI-----TQGR--VLQRPRTCPQEVYELMLGCGWQREPHMRKNIKGI 828

QY 902 HRFL 905
DB 829 HTLL 832

RESULT 11
US-08-359-705B-2
Sequence 2, Application US/08359705B
Patent No. 5844092
GENERAL INFORMATION:
APPLICANT: Presta, Leonard G.
APPLICANT: Shelton, David L.
APPLICANT: Urfer, Roman
TITLE OF INVENTION: Human trk Receptors and Neurotrophic Factor Inhibitors
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/359,705B
FILING DATE: 20-Dec-1994
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/286846
```

```
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 854C1P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 847 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-286-305A-5

Query Match 13.4%; Score 658; DB 2; Length 847;
Best Local Similarity 47.4%; Pred. No. 6.9e-40;
Matches 144; Conservative 45; Mismatches 79; Indels 36; Gaps 9;

QY 607 RSLRFEKELGEGQGEVHLCE---VDSPODLVSLDFPLNVRKGHPLLVAVKILRPDATK 663
DB 560 RHNIYVKRELGEAGFGKVFLEACYNLCPEQD-----KILVAVKTLK-DASD 604

QY 664 NARNDLFKEVKIMSLKDPNIIRLLGVCVODDPLCMITDYMENGLNOFLSAHQLED-KA 722
DB 605 NARKDFHREAELLTNLQHEHIVKFGVCVEGDPPLIMVFYMKHGLNKLFLRAHGPDAVLM 664

QY 723 AEGAPGDGQAAQGPT-ISYPMLLHVAQAISGMRYLATLNFVHRDLATRNCLVGENFTIK 781
DB 665 AEGNP-----PTELTSQMLHIAQIAAGMVLASQHFVHRDLATRNCLVGENLLVK 716

QY 782 IADFGMSRLNLYAGDYRYVQGRAVLPKRWMAWECILMGKFTTASDWAFGVTLWEVLMCLR 841
DB 717 IGDFGMSRDYSTDYRYVGGHTMLPIRWMPPEIMRYKFTTESDWLSGVVLEIFYGK 776

QY 842 AQPFGOLTDEQVIENAGEFFRDQGRQVYLSRPPACPGQLYELMLRCWSRESEQRPPFSOL 901
DB 777 -QPWYQLSNNVEIECI-----TQGR--VLQRPRTCPQEVYELMLGCGWQREPHMRKNIKGI 828

QY 902 HRFL 905
DB 829 HTLL 832

RESULT 11
US-08-359-705B-2
Sequence 2, Application US/08359705B
Patent No. 5844092
GENERAL INFORMATION:
APPLICANT: Presta, Leonard G.
APPLICANT: Shelton, David L.
APPLICANT: Urfer, Roman
TITLE OF INVENTION: Human trk Receptors and Neurotrophic Factor Inhibitors
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/359,705B
FILING DATE: 20-Dec-1994
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/286846
```

```
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 00,000
REFERENCE/DOCKET NUMBER: 854C1P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 847 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-286-305A-5

Query Match 13.4%; Score 658; DB 2; Length 847;
Best Local Similarity 47.4%; Pred. No. 6.9e-40;
Matches 144; Conservative 45; Mismatches 79; Indels 36; Gaps 9;

QY 607 RSLRFEKELGEGQGEVHLCE---VDSPODLVSLDFPLNVRKGHPLLVAVKILRPDATK 663
DB 560 RHNIYVKRELGEAGFGKVFLEACYNLCPEQD-----KILVAVKTLK-DASD 604

QY 664 NARNDLFKEVKIMSLKDPNIIRLLGVCVODDPLCMITDYMENGLNOFLSAHQLED-KA 722
DB 605 NARKDFHREAELLTNLQHEHIVKFGVCVEGDPPLIMVFYMKHGLNKLFLRAHGPDAVLM 664

QY 723 AEGAPGDGQAAQGPT-ISYPMLLHVAQAISGMRYLATLNFVHRDLATRNCLVGENFTIK 781
DB 665 AEGNP-----PTELTSQMLHIAQIAAGMVLASQHFVHRDLATRNCLVGENLLVK 716

QY 782 IADFGMSRLNLYAGDYRYVQGRAVLPKRWMAWECILMGKFTTASDWAFGVTLWEVLMCLR 841
DB 717 IGDFGMSRDYSTDYRYVGGHTMLPIRWMPPEIMRYKFTTESDWLSGVVLEIFYGK 776

QY 842 AQPFGOLTDEQVIENAGEFFRDQGRQVYLSRPPACPGQLYELMLRCWSRESEQRPPFSOL 901
DB 777 -QPWYQLSNNVEIECI-----TQGR--VLQRPRTCPQEVYELMLGCGWQREPHMRKNIKGI 828

QY 902 HRFL 905
DB 829 HTLL 832

RESULT 11
US-08-359-705B-2
Sequence 2, Application US/08359705B
Patent No. 5844092
GENERAL INFORMATION:
APPLICANT: Presta, Leonard G.
APPLICANT: Shelton, David L.
APPLICANT: Urfer, Roman
TITLE OF INVENTION: Human trk Receptors and Neurotrophic Factor Inhibitors
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/359,705B
FILING DATE: 20-Dec-1994
CLASSIFICATION: 424
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 08/286846
```

```

; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, PhD., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P0873P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-8674
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 822 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-08-286-846A-2

Query Match          13.4%; Score 658; DB 2; Length 822;
Best Local Similarity 47.4%; Pred. No. 6.6e-40;
Matches 144; Conservative 45; Mismatches 79; Indels 36; Gaps 9;

QY 607 RSLRFEKELGEGQFGEVHLCE---VDSQDLVSLDFPLNVRKGHPLLVAVKILRPDATK 663
Db 535 RHNVILKRELGEAGFKVFLAECYNLCPQD-----KILVAVKTLK-DASD 579

QY 664 NARNDFLKEVKIMSLKDPNIIRLLGVCVQDDPLCMITDYMENGDLNQLSAHQLED-KA 722
Db 580 NARKDFHREAELTNLQHEHIVKIFYGVCVEGDPLIMVFEYMKHGDNLKFLRAHGPDAVLM 639

QY 723 AEGAPDGGQAAGQPT-ISYPMLLHVAQAQASGMRYLATLNFVHRDLATNCLVGENFTIK 781
Db 640 AEGNP-----PTELTSQMLHIAQAQTAAGNVIYLAHQFVHRDLATNCLVGENLLVK 691

QY 782 IADFGMSRNLVAGDYRVQGRAVLPPIRMMAWECILMGKFTTASDVWAFGVTLWEVLMCLR 841
Db 692 IGDGMSRDVYSTDYRVGGHTMLPIRMWPPESIMYRKFTTESDVWLSGVWLWEIFTYCK 751

QY 842 AQPFGQLTDEQVIENAGEFRDQGRQVYLSRPPACQGLYELMLRCWSRESQRPFFSOL 901
Db 752 -QPWYQLSNNEVIECI-----TQGR--VLORPRTCPQEVYELMLRCWSQREPHMRKNIKGI 803

QY 902 HRFL 905
Db 804 HTLL 807

RESULT 13
US-08-441-104A-5
; Sequence 5, Application US/08441104A
; Patent No. 5891650
; GENERAL INFORMATION:
; APPLICANT: Godowski, Paul J.
; APPLICANT: Mark, Melanie R.
; APPLICANT: Sadick, Michael D.
; APPLICANT: Shelton, David L.
; APPLICANT: Wong, Wai Lee Tan
; TITLE OF INVENTION: KINASE RECEPTOR ACTIVATION ASSAY
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 720 kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/441.104A
; FILING DATE: 15-MAY-1995

```

```

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/286305
; FILING DATE: 05-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/170558
; FILING DATE: 20-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/157563
; FILING DATE: 23-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 00,000
; REFERENCE/DOCKET NUMBER: 854C1PIC2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 847 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-441-104A-5

Query Match          13.4%; Score 658; DB 2; Length 847;
Best Local Similarity 47.4%; Pred. No. 6.9e-40;
Matches 144; Conservative 45; Mismatches 79; Indels 36; Gaps 9;

QY 607 RSLRFEKELGEGQFGEVHLCE---VDSQDLVSLDFPLNVRKGHPLLVAVKILRPDATK 663
Db 560 RHNVILKRELGEAGFKVFLAECYNLCPQD-----KILVAVKTLK-DASD 604

QY 664 NARNDFLKEVKIMSLKDPNIIRLLGVCVQDDPLCMITDYMENGDLNQLSAHQLED-KA 722
Db 605 NARKDFHREAELTNLQHEHIVKIFYGVCVEGDPLIMVFEYMKHGDNLKFLRAHGPDAVLM 664

QY 723 AEGAPDGGQAAGQPT-ISYPMLLHVAQAQASGMRYLATLNFVHRDLATNCLVGENFTIK 781
Db 665 AEGNP-----PTELTSQMLHIAQAQTAAGNVIYLAHQFVHRDLATNCLVGENLLVK 716

QY 782 IADFGMSRNLVAGDYRVQGRAVLPPIRMMAWECILMGKFTTASDVWAFGVTLWEVLMCLR 841
Db 717 IGDGMSRDVYSTDYRVGGHTMLPIRMWPPESIMYRKFTTESDVWLSGVWLWEIFTYCK 776

QY 842 AQPFGQLTDEQVIENAGEFRDQGRQVYLSRPPACQGLYELMLRCWSRESQRPFFSOL 901
Db 777 -QPWYQLSNNEVIECI-----TQGR--VLORPRTCPQEVYELMLRCWSQREPHMRKNIKGI 828

QY 902 HRFL 905
Db 829 HTLL 832

RESULT 14
US-08-457-880A-2
; Sequence 2, Application US/08457880A
; Patent No. 5910574
; GENERAL INFORMATION:
; APPLICANT: Leonard G. Presta
; APPLICANT: David L. Shelton
; APPLICANT: Roman Urfer
; TITLE OF INVENTION: HUMAN trk RECEPTORS AND NEUROTROPHIC FACTOR
; NUMBER OF SEQUENCES: 41
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080

```

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/457,880A  
FILING DATE:  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/444,622  
FILING DATE: 19-May-1995  
APPLICATION NUMBER: 08/286846  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Torchia, PhD., Timothy E.  
REGISTRATION NUMBER: 36,700  
REFERENCE/DOCKET NUMBER: P0873PIC3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-8674  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 822 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-08-457-880A-2

Query Match 13.4%; Score 658; DB 2; Length 822;  
Best Local Similarity 47.4%; Pred. No. 6.6e-40;  
Matches 144; Conservative 45; Mismatches 79; Indels 36; Gaps 9;

QY 607 RSRLEKELGEGGEGVHLC-----VDSQDLVSLDFPLNVRKGHPDLLVAVKILRPD ATK 663  
DB 535 RHNVILKRELGEAGFKVFLAECYNLCPEQD-----KILVAVKTLK-DASD 579  
QY 664 NARNDFLKEVKIMSRKDPNIIRLLGVCVQDDPLCMITDYMGDNLQFSAHQLED-KA 722  
DB 580 NARKDFHREAEELLNLQHEHIVKFGVCGVEGDLIMVFEYMKHGDNLKFLRAHGPDAVLM 639  
QY 723 AEGAPGDQQAAGGT-ISYPMLLHVAQAISGMRYLATLNFVHRDLATRNCLVGENFTIK 781  
DB 640 AEGNP-----PTLTSQMLHQAQIAAGMVYLAHQFVHRDLATRNCLVGENLLVK 691  
QY 782 IADFGMSRLNLYAGDYRVQGRAVLPIRMWAEICILMGKFTTASDVWAFGYTLWEVLMCLR 841  
DB 692 IGDFGMSRDVSTDYRYVGGHTMLPIRMWPEIMYRKFTTESDVWSLGVVLEIFTYK 751  
QY 842 AQPFGQLTDEQVIENAGEFFRDQGRQVYLSRPPACPGQLYELMLRCWSRESEORPPFSOL 901  
DB 752 -QPWYQLSNNEVIECI-----TQGR--VLQRPRTCPQEVYELMLGCGWQREPHMRKNIKGI 803  
QY 902 HRFL 905  
DB 804 HTLL 807

RESULT 15  
US-08-440-816A-5  
Sequence 5, Application US/08440816A  
Patent No. 5914237  
GENERAL INFORMATION:  
APPLICANT: Godowski, Paul J.  
APPLICANT: Mark, Melanie R.  
APPLICANT: Sadick, Michael D.  
APPLICANT: Shelton, David L.  
APPLICANT: Wong, Wai Lee Tan  
TITLE OF INVENTION: KINASE RECEPTOR ACTIVATION ASSAY  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd

CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/440,816A  
FILING DATE: 15-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/286305  
FILING DATE: 05-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/170558  
FILING DATE: 20-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/157563  
FILING DATE: 23-NOV-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Lee, Wendy M.  
REGISTRATION NUMBER: 00,000  
REFERENCE/DOCKET NUMBER: 854CIP1C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-1994  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 847 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-440-816A-5

Query Match 13.4%; Score 658; DB 2; Length 847;  
Best Local Similarity 47.4%; Pred. No. 6.9e-40;  
Matches 144; Conservative 45; Mismatches 79; Indels 36; Gaps 9;

QY 607 RSRLEKELGEGGEGVHLC-----VDSQDLVSLDFPLNVRKGHPDLLVAVKILRPD ATK 663  
DB 560 RHNVILKRELGEAGFKVFLAECYNLCPEQD-----KILVAVKTLK-DASD 604  
QY 664 NARNDFLKEVKIMSRKDPNIIRLLGVCVQDDPLCMITDYMGDNLQFSAHQLED-KA 722  
DB 605 NARKDFHREAEELLNLQHEHIVKFGVCGVEGDLIMVFEYMKHGDNLKFLRAHGPDAVLM 664  
QY 723 AEGAPGDQQAAGGT-ISYPMLLHVAQAISGMRYLATLNFVHRDLATRNCLVGENFTIK 781  
DB 665 AEGNP-----PTLTSQMLHQAQIAAGMVYLAHQFVHRDLATRNCLVGENLLVK 716  
QY 782 IADFGMSRLNLYAGDYRVQGRAVLPIRMWAEICILMGKFTTASDVWAFGYTLWEVLMCLR 841  
DB 717 IGDFGMSRDVSTDYRYVGGHTMLPIRMWPEIMYRKFTTESDVWSLGVVLEIFTYK 776  
QY 842 AQPFGQLTDEQVIENAGEFFRDQGRQVYLSRPPACPGQLYELMLRCWSRESEORPPFSOL 901  
DB 777 -QPWYQLSNNEVIECI-----TQGR--VLQRPRTCPQEVYELMLGCGWQREPHMRKNIKGI 828  
QY 902 HRFL 905  
DB 829 HTLL 832

Search completed: November 4, 1999, 01:52:00  
Job time: 7590 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 3, 1999, 13:36:38 ; Search time 22.34 Seconds  
(without alignments)  
968.013 Million cell updates/sec

Title: US-08-170-558-4  
Perfect score: 4903  
Sequence: 1 MGPEALSSLLLLLVASGDA.....QRPFSQHLRFLAEDALNTV 913

Scoring table: BLOSUM62

Searched: 188963 seqs, 23686106 residues

Database: A\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	4882	99.6	919	1	W34672 Human mammary carc
2	4875	99.4	919	1	W75502 Human mammary carc
3	4875	99.4	919	1	W75504 Human mammary carc
4	4852.5	99.0	914	1	R71100 Protein-tyrosine-k
5	4864.5	95.1	876	1	W34675 Human mammary carc
6	4651.5	94.9	882	1	W34673 Human mammary carc
7	3011	61.4	563	1	W34674 Human mammary carc
8	2415	49.3	855	1	W75503 Human colonic aden
9	2415	49.3	855	1	W75505 Human colonic aden
10	2415	49.3	855	1	W34671 CCK-2, a human mam
11	2415	49.3	855	1	W71114 Discoidin domain r
12	2413	49.2	854	1	W79152 Receptor protein t
13	2413	49.2	854	1	W81409 Receptor protein t
14	1825	37.2	650	1	R54089 Partial sequence o
15	658	13.4	822	1	R81630 Human trkB recepto
16	658	13.4	847	1	W11941 qD.trkB fusion use
17	650.5	13.3	825	1	R81631 Human trkB recepto
18	650.5	13.3	850	1	W11942 qD.trkB fusion use
19	642.5	13.1	825	1	R62021 Porcine trkB prote
20	639.5	13.0	825	1	R27148 Adult porcine trkB
21	639.5	13.0	825	1	R30883 trkB gene prod. Nu
22	638	13.0	814	1	W11940 qD.trkB fusion use
23	637	13.0	830	1	R81627 Human trkB recepto
24	633.5	12.9	839	1	R81625 Human trkB recepto
25	627	12.8	728	1	R30884 Mouse trkB proto-o
26	627	12.8	728	1	R30884 Partial trkB gene
27	627	12.8	728	1	R62022 Murine trkB protei
28	625.5	12.8	839	1	R71621 Porcine trkB K2 is
29	621.5	12.7	739	1	R71618 Murine trkB K3 iso
30	606.5	12.4	868	1	W26507 Rat Dmk receptor.
31	606.5	12.4	868	1	W26610 Rat muscle-specifi
32	605.5	12.3	881	1	R84091 Nsk2 receptor with
33	603	12.3	868	1	R92717 Mouse muscle-local
34	603	12.3	871	1	R84087 Nsk2 receptor. Nuc
35	603	12.3	871	1	W62568 Mouse receptor tyr
36	603	12.3	867	1	W62583 Mouse receptor tyr
37	601.5	12.3	869	1	W26506 Human Dmk receptor
38	601.5	12.3	869	1	W26611 Human muscle-speci
39	601.5	12.3	861	1	W62572 Mouse Nsk2 (altern
40	601	12.3	873	1	R84092 Nsk2 receptor with
41	600.5	12.2	530	1	R92715 Mouse muscle-local
42	600.5	12.2	860	1	R92716 Mouse muscle-local
43	600.5	12.2	354	1	R94989 Nsk2 receptor intr

## ALIGNMENTS

RESULT	ID	Score	Length	DB ID	Description
1	W34672	44	600.5	12.2	863 1 R84088
	W34672	45	600.5	12.2	863 1 W62569
	W34672				standard; Protein; 919 AA.
	W34672				(first entry)
	W34672				Human mammary carcinoma kinase 10 (MCK-10) amino acid sequence.
	W34672				Mammary carcinoma kinase; MCK-10; receptor tyrosine kinase;
	W34672				proliferative disease; cancer; insulin receptor family;
	W34672				tyrosine kinase neurotrophin receptor; MCK-10 activity;
	W34672				neurological disorder; aberrant expression.
	W34672				Homo sapiens.
	W34672				Location/Qualifiers
	W34672				1..18
	W34672				Peptide
	W34672				/label= signal_sequence
	W34672				19..919
	W34672				Protein
	W34672				/note= "mature_protein"
	W34672				31..185
	W34672				Domain
	W34672				/label= Discoidin_I_like_domain
	W34672				304..307
	W34672				Cleavage_site
	W34672				/label= endopeptidase_furin
	W34672				/note= "putative precursor cleavage site"
	W34672				48..439
	W34672				Region
	W34672				/label= transmembrane_region
	W34672				617..627
	W34672				Binding_site
	W34672				/label= ATP_binding_motif
	W34672				802..803
	W34672				Modified_site
	W34672				/label= autophosphorylation_sites
	W34672				/note= "putative"
	W34672				798
	W34672				Modified_site
	W34672				/label= autophosphorylation_site
	W34672				/note= "putative"
	W34672				844..847
	W34672				Binding_site
	W34672				/label= binding_motif_for_P13_kinase
	W34672				/note= "binding motif for phosphatidylinositol 3'
	W34672				kinase"
	W34672				832..832
	W34672				Binding_site
	W34672				/label= potential_substrate_binding_site
	W34672				506..509
	W34672				Binding_site
	W34672				/label= putative_receptor_binding_site_for_SHC
	W34672				/note= "SHC is an oncogenic SH2 domain containing molecule"
	W34672				510..513
	W34672				Binding_site
	W34672				/label= GTPase_activity_protein_binding_site
	W34672				/note= "putative"
	W34672				505..541
	W34672				Region
	W34672				/note= "alternatively spliced variant"
	W34672				666..671
	W34672				Region
	W34672				/note= "alternatively spliced sequence"
	W34672				26..42
	W34672				Region
	W34672				/note= "antibody recognition sequence NTalpha"
	W34672				309..321
	W34672				Region
	W34672				/note= "antibody recognition sequence NTbeta"
	W34672				902..919
	W34672				Region
	W34672				/note= "antibody recognition sequence CTbeta"

US5677144-A.  
14-OCT-1997.  
08-NOV-1994; 336343.  
16-NOV-1993; US-153397.  
(ALVE)/ ALVES F H E.  
(ULLR)/ ULLRICH A.  
Alves FHE, Ullrich A;  
WPI: 97-511869/47.  
N-FSD; 793785.  
Truncated receptor tyrosine kinase CCK-2 - and nucleic acid coding  
for it, useful for cancer diagnosis  
Disclosure; Fig 1; 70pp; English.

The present sequence represents the protein sequence of a mammary carcinoma kinase, called MCK-10. This kinase belongs to a novel family of receptor tyrosine kinases, and expression is associated with proliferative diseases such as cancer. The MCK-10 receptor tyrosine kinase has extensive sequence similarity to the insulin receptor family. The MCK-10 gene was isolated by PCR using 2 degenerate oligonucleotide primer pools, using a template cDNA synthesised by reverse transcription of poly-A RNA from the human mammary carcinoma cell line MCF7. The MCK-10 protein contains 2 alternative spliced sequences, from amino acids 505-541 and 666-671. The sequence represented by amino acids 585-595 may be important, as deletion of this motif in the activin receptor serine/threonine kinase results in reduced ligand binding affinity. MCK-10 is expressed in brain tissue, and the protein shares homology with the tyrosine kinase neurotrophin receptor. Modulation of MCK-10 activity therefore may be used for treatment of neurological disorders. MCK-10 is also expressed in a variety of cancer cell lines and tumour tissue. The nucleotide sequence of MCK-10, or parts of it, can be used for diagnostic purposes to detect aberrant expression of MCK-10 genes. Inhibitors of MCK-10 receptor activity may have therapeutic value in the treatment of diseases such as cancer.

Sequence 919 AA:

Query Match 99.6% Score 4882; DB 1; Length 919;  
 Best Local Similarity 99.2%; Pred. No. 0;  
 Matches 912; Conservative 0; Mismatches 1; Indels 6; Gaps 1;

QY 1 MGPEALSLLLVLSAGDADMKHFDPAKRYALGMDRTIPDSDISASSWSDDTAAR 60  
 DB 1 MGPEALSLLLVLSAGDADMKHFDPAKRYALGMDRTIPDSDISASSWSDDTAAR 60

QY 61 HSRLESSDGDGAWCPAGSVFPKKEEYLQVDLQRLHLVALVGTQGRHAGGLGKFFSRYL 120  
 DB 61 HSRLESSDGDGAWCPAGSVFPKKEEYLQVDLQRLHLVALVGTQGRHAGGLGKFFSRYL 120

QY 121 RYSDGRRWGCKDRWGQEVISGNEDEGVVYLDLGPMPVARLVFPYPRADRVMSVCLRV 180  
 DB 121 RYSDGRRWGCKDRWGQEVISGNEDEGVVYLDLGPMPVARLVFPYPRADRVMSVCLRV 180

QY 181 ELYGCLWRDGLSYTAPVQGTMYLSEAVYLNDSYDGTHTVGGYGLGQLADGVVGLDD 240  
 DB 181 ELYGCLWRDGLSYTAPVQGTMYLSEAVYLNDSYDGTHTVGGYGLGQLADGVVGLDD 240

QY 241 FRKSOELRVPGYDYGVGNSHNSFSGVEVEFEFDRLRARAFQAMQVHCNNMHTLGARLPGG 300  
 DB 241 FRKSOELRVPGYDYGVGNSHNSFSGVEVEFEFDRLRARAFQAMQVHCNNMHTLGARLPGG 300

QY 301 VECFRGCPAMAWEGEPMRHNLGNLGDPRARAVSVPLGGRVARFLQCRFLFAGPWLIFS 360  
 DB 301 VECFRGCPAMAWEGEPMRHNLGNLGDPRARAVSVPLGGRVARFLQCRFLFAGPWLIFS 360

QY 361 EISFISDVVNNSSPALGGTTPPAPWPPGPPPTNFSSLELEPRGQQPVAKPEGSPTAILI 420  
 DB 361 EISFISDVVNNSSPALGGTTPPAPWPPGPPPTNFSSLELEPRGQQPVAKPEGSPTAILI 420

QY 421 GCLVAITLILLIITALLMLRWLRHLWRLLSKARRVLEELIVHLSVPGDITLINNRGPRE 480  
 DB 421 GCLVAITLILLIITALLMLRWLRHLWRLLSKARRVLEELIVHLSVPGDITLINNRGPRE 480

QY 481 PPYQEPFRPGNPPHSPAPVNGSALLSNPARYALLATYARPPRGPPPTPAWAKPTNT 540  
 DB 481 PPYQEPFRPGNPPHSPAPVNGSALLSNPARYALLATYARPPRGPPPTPAWAKPTNT 540

QY 541 QAYSGDYMEPEKPGAPLLPPPPQNSVPHYAEADIVTLQGVGTGNTYAVPALPPGAVGDGP 600  
 DB 541 QAYSGDYMEPEKPGAPLLPPPPQNSVPHYAEADIVTLQGVGTGNTYAVPALPPGAVGDGP 600

QY 601 PRVDFPSRLRFKLEKGEQFGEVHLCVDSQDLSLDFPLNVRKGPLLVAVKILRPD 660  
 DB 601 PRVDFPSRLRFKLEKGEQFGEVHLCVDSQDLSLDFPLNVRKGPLLVAVKILRPD 660

QY 661 ATKNA-----RNDFLKEVKIMSRLLKDPNIIRLGVCVQDDPLCMITDYMENGDLNQELS 714  
 DB 661 ATKNA-----RNDFLKEVKIMSRLLKDPNIIRLGVCVQDDPLCMITDYMENGDLNQELS 714

Db 661 ATKNASFSLFSRNDLKEVKIMSRLLKDPNIIRLGVCVQDDPLCMITDYMENGDLNQELS 720  
 QY 715 AHQLEDKRAEGAPGQQAAGQPTISYPMLLHVAAQIASGMRYLATLNFVHRDLATRNCLV 774  
 Db 721 AHQLEDKRAEGAPGQQAAGQPTISYPMLLHVAAQIASGMRYLATLNFVHRDLATRNCLV 780  
 QY 775 GENFTIKIADFGMSRNLYAGDYRYVQGRAVLPIRMAWECILMGKFTTASDVWAFGVTLW 834  
 Db 781 GENFTIKIADFGMSRNLYAGDYRYVQGRAVLPIRMAWECILMGKFTTASDVWAFGVTLW 840  
 QY 835 EYVLMCLRAQPEGQLTDEQVIENAGEFFRDQGRQVYLSRPPACPOGLYELMLRCWSRESEQ 894  
 Db 841 EYVLMCLRAQPEGQLTDEQVIENAGEFFRDQGRQVYLSRPPACPOGLYELMLRCWSRESEQ 900  
 QY 895 RPPFSQLHRFLAEDALNTV 913  
 Db 901 RPPFSQLHRFLAEDALNTV 919

RESULT 2  
 R75502  
 ID R75502 standard; Protein; 919 AA.  
 AC R75502;  
 DT 26-NOV-1995 (first entry)  
 DE Human mammary carcinoma kinase 10 (MCK-10).  
 KW Mammary carcinoma kinase 10; transmembrane receptor;  
 KW receptor tyrosine kinase; cancer.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT peptide 1..18  
 FT domain /label= signal  
 FT 31..185 /label= discoidin I-like domain  
 FT cleavage\_site 304..307 /label= putative precursor cleavage site  
 FT region 417..439 /label= transmembrane  
 FT misc\_difference 505..541 /label= alternatively spliced sequence I  
 FT misc\_difference 666..671 /label= alternatively spliced sequence II  
 FT misc\_difference 25..42 /label= NT alpha  
 FT /note= "peptide antibody recognition site"  
 FT misc\_difference 309..321 /label= NT beta  
 FT /label= CT beta  
 FT misc\_difference 909..919 /note= "see above"  
 FT /label= CT beta  
 FT /note= "see above"  
 PN W09514088-A.  
 PD 26-MAY-1995.  
 PF 16-NOV-1994; E03797.  
 PR 16-NOV-1993; US-153397.  
 PA (PLAC ) MAX PLANCK GES. FOERDERUNG WISSENSCHAFTEN.  
 PI Alves FHE, Ullrich A;  
 DR WPI; 95-224054/29.  
 DR N-PSDB; Q92520.  
 PT New nucleic acid encoding MCK-10 receptor tyrosine kinase - and  
 PT derived vectors, transformed cells, proteins and antibodies useful  
 PT for diagnosis and treatment of proliferative disease, esp. cancer,  
 PT and for screening modulators  
 PS Disclosure: Page 53-55; 115pp; English.  
 CC cDNA prep. from human breast cancer cell line MCF7 (ATCC HTB22) and  
 CC used in a PCR with two degenerate oligo primer pools based on  
 CC conserved sequences of the kinase domain of receptor tyrosine  
 CC kinases. One clone, designated MCK-10, was identified as novel RTK.  
 CC The PCR fragment was used to screen a lambda gt11 library of human  
 CC fetal brain cDNA. Several overlapping clones were identified. The  
 CC composite of these cDNA clones is given in Q92520 and the deduced AA  
 CC sequence in R75502. Some of the clones had a deletion of 6 AAs at  
 CC posn. 2315 in the MCK-10 sequence. MCK-10 has all the  
 CC characteristics of a receptor PK (see R75502 FT). Screening of



CC human placental library yielded two cDNA clones MCK-10-1 and  
 CC MCK-10-2. One of the clones isolated from the human fetal brain  
 CC library contd. an additional 18 nts in the TK domain. The MCK-10 splice  
 CC isoforms have been designated MCK-10-1 (with an additional 111 bp between  
 CC nts 1832 and 1943); MCK-10-2 (without any insertions); MCK-10-3 (with the  
 CC additional 111 bp and 18 bp in the TK domain); and MCK-10-4 (with the  
 CC additional 18 bp). The predicted mol. wts. of MCK-10-1 and MCK-10-2  
 CC pre-receptors are 101.13 and 97.17 kD respectively, and can thus be  
 CC subdivided into a 34.31 kD alpha subunit and a 66.84 or 62.88 kD  
 CC beta subunits that contain the TK homology and alternative splice sites.  
 SQ Sequence 919 AA;

Query Match 99.4%; Score 4875; DB 1; Length 919;  
 Best Local Similarity 99.1%; Pred. NO. 0;  
 Matches 911; Conservative 0; Mismatches 2; Indels 6; Gaps 1;

QY 1 MGPEALSSLLLLVAGSDADMKHFDPAKCRYALGMDRTTIPDSISASSWSDDSTAAR 60  
 DB 1 MGPEALSSLLLLVAGSDADMKHFDPAKCRYALGMDRTTIPDSISASSWSDDSTAAR 60

QY 61 HSRLESSDGDGAWCPAGSVFPKEEYLVQDLQRLHLVALVGTQGRHAGGLGKEFSRSLR 120  
 DB 61 HSRLESSDGDGAWCPAGSVFPKEEYLVQDLQRLHLVALVGTQGRHAGGLGKEFSRSLR 120

QY 121 RYSDGRRWGWKDRGQEVISGNEDEPGVYLKDLGPPWVARLRYFPYPRADRVMSVCLRV 180  
 DB 121 RYSDGRRWGWKDRGQEVISGNEDEPGVYLKDLGPPWVARLRYFPYPRADRVMSVCLRV 180

QY 181 ELYGCLWRDGLLSYAPVQOTMYLSEAVYLVNDSTVDGHTVGLQYGGGLQADGVVGLDD 240  
 DB 181 ELYGCLWRDGLLSYAPVQOTMYLSEAVYLVNDSTVDGHTVGLQYGGGLQADGVVGLDD 240

QY 241 FRKSOELRWPCGYDYGVGNSHFSFGYVEMEFEDRLRAFAQMVHCHNMHTLGARLPGG 300  
 DB 241 FRKSOELRWPCGYDYGVGNSHFSFGYVEMEFEDRLRAFAQMVHCHNMHTLGARLPGG 300

QY 301 VECFRFRGPAMAWEGEPHRLNGLGNDPRARAVSVPLGGRVAREPLOCPLFAGPWLIFS 360  
 DB 301 VECFRFRGPAMAWEGEPHRLNGLGNDPRARAVSVPLGGRVAREPLOCPLFAGPWLIFS 360

QY 361 EISFISDVVNNSSPALGGTFPPAPWPPGPTTFNFSLELEPRGQOPVAKPEGSPPTALLI 420  
 DB 361 EISFISDVVNNSSPALGGTFPPAPWPPGPTTFNFSLELEPRGQOPVAKPEGSPPTALLI 420

QY 421 GCLVAIIILLLLIIALMLRWLHWRLLSKAERRVLEELTVHLSVPGDTILINNRPGPRE 480  
 DB 421 GCLVAIIILLLLIIALMLRWLHWRLLSKAERRVLEELTVHLSVPGDTILINNRPGPRE 480

QY 481 PPYQEPFRGNPPHSPACVNGSALLLSNPAYRLLLATYARPPRGPGPPTPAWAKPTNT 540  
 DB 481 PPYQEPFRGNPPHSPACVNGSALLLSNPAYRLLLATYARPPRGPGPPTPAWAKPTNT 540

QY 541 QAYSQDYMEPEKPGAPLLPPPPQNSVPHYADIVTLQGVGTGNTYAVPALPPCAVGDGP 600  
 DB 541 QAYSQDYMEPEKPGAPLLPPPPQNSVPHYADIVTLQGVGTGNTYAVPALPPCAVGDGP 600

QY 601 PRVDFPRSLRFEKELGSGQGEVHLCVDSQDVLVSDFPLNVRKGHPLLVAVKILRPD 660  
 DB 601 PRVDFPRSLRFEKELGSGQGEVHLCVDSQDVLVSDFPLNVRKGHPLLVAVKILRPD 660

QY 661 ATKNA-----RNDPLKEVKIMSLKDPNIIIRLLGVCVQDDPLCMITDYMGNDLQFLS 714  
 DB 661 ATKNAFSLFRNDPLKEVKIMSLKDPNIIIRLLGVCVQDDPLCMITDYMGNDLQFLS 714

QY 715 AHQLEDKAAEGAPGDQAAGQPTTISYPMLLHVAQAQIASGMRYLATLNFVHRDLATRNCLV 774  
 DB 721 AHQLEDKAAEGAPGDQAAGQPTTISYPMLLHVAQAQIASGMRYLATLNFVHRDLATRNCLV 780

QY 775 GENFTIKIADFGMSRLNLAGDYRYRQGRAVLPPIRMMAWECILMGKFTTASDVWAFGVTW 834  
 DB 781 GENFTIKIADFGMSRLNLAGDYRYRQGRAVLPPIRMMAWECILMGKFTTASDVWAFGVTW 840

QY 835 EVLMCLRAQPFQGLTDEQVIENAGEFRDQGRVYLRRPACQGLYELMLRCWSRESEQ 894  
 DB 841 EVLMCLRAQPFQGLTDEQVIENAGEFRDQGRVYLRRPACQGLYELMLRCWSRESEQ 900

QY 895 RPPFSQLHRLAEDALNTV 913  
 DB 901 RPPFSQLHRLAEDALNTV 919

RESULT 3  
 R75504  
 ID R75504 standard; Protein; 919 AA.  
 AC R75504;  
 DT 26-NOV-1995 (first entry)  
 DE Human mammary carcinoma kinase 10 (MCK-10).  
 KW Mammary carcinoma kinase 10; MCK-10; transmembrane receptor;  
 KW receptor tyrosine kinase; cancer.  
 OS Homo sapiens.  
 FH Key  
 FT peptide  
 FT 1..18 Location/Qualifiers  
 FT 31..185 /label= signal  
 FT domain /label= discooidin I-like domain  
 FT cleavage\_site 304..307  
 FT region 417..439 /label= putative precursor cleavage site  
 FT misc\_difference 505..541 /label= transmembrane  
 FT misc\_difference 666..671 /label= alternatively spliced sequence I  
 FT misc\_difference 25..42 /label= alternatively spliced sequence II  
 FT /label= NT alpha  
 FT /note= "peptide antibody recognition site"  
 FT misc\_difference 309..331 /label= NT beta  
 FT /note= "see above"  
 FT misc\_difference 909..919 /label= CT beta  
 FT /note= "see above"  
 PN W09514089-A.  
 PD 26-MAY-1995.  
 PF 16-NOV-1994; E03799.  
 PR 16-NOV-1993; US-153397.  
 PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.  
 PI Alves FHE, Ullrich A;  
 DR WPI: 95-224055/29.  
 DR N-PSDB: Q92522.  
 PT New nucleic acid encoding CCK-2 receptor tyrosine kinase - and  
 PT derived vectors, transformed cells, proteins and antibodies, useful  
 PT for diagnosis and treatment of proliferative and nervous system  
 PT diseases and for screening modulators  
 PS Disclosure: Page 70-72; 115pp; English.  
 CC CDNA prepd. from human breast cancer cell line MCF7 (ATCC HTB22) was  
 CC used in a PCR with two degenerate oligo primer pools based on  
 CC conserved sequences of the kinase domain of receptor tyrosine  
 CC kinases. One clone, designated MCK-10, was identified as novel RTK.  
 CC The PCR fragment was used to screen a lambda gt11 library of human  
 CC fetal brain cDNA. Several overlapping clones were identified. The  
 CC composite of these cDNA clones is given in Q92522 and the deduced AA  
 CC sequence in R75504. Some of the clones had a deletion of 6AA at posn.  
 CC 2315 in the MCK-10 sequence. MCK-10 has all the characteristics of  
 CC a receptor PTK (see R75504 FT). Screening of human placental library  
 CC yielded two cDNA clones. One of the clones isolated from the human  
 CC fetal brain library contained an additional 18 nts in the TK  
 CC domain. The MCK-10 splice isoforms have been designated MCK-10-1  
 CC (with an additional 111 bp between nts 1832 and 1943); MCK-10-2  
 CC (without any insertions); MCK-10-3 (with the additional 111 bps and  
 CC 18 bp in the TK domain); and MCK-10-4 (with the additional 18 bp).  
 CC The predicted mol. wts. of MCK-10-1 and MCK-10-2 pre-receptors are  
 CC 101.13 and 97.17 kD respectively, and can thus be subdivided into a  
 CC 34.31 kD alpha subunit and a 66.84 or 62.88 kD beta subunits that  
 CC contain the TK homology and alternative splice sites.

SQ Sequence 919 AA;

Query Match 99.48; Score 4875; DB 1; Length 919;  
Best Local Similarity 99.18; Pred. No. 0;

Qy	1	MGPEALSLLLLLVASGDADKMGHPDPKAKRYALGMDQRTTIPDSDI	SASSWSWSDSTAAR	60
Db	1	MGPEALSLLLLLVASGDADKMGHPDPKAKRYALGMDQRTTIPDSDI	SASSWSWSDSTAAR	60
Qy	61	HSRLESSDGDGAWCPAGSVFPKEEYLYQVDLQRLHLVALVGTQGRHAGGLGKEFSRSYRL	120	
Db	61	HSRLESSDGDGAWCPAGSVFPKEEYLYQVDLQRLHLVALVGTQGRHAGGLGKEFSRSYRL	120	
Qy	121	RYSRDGRRWGMKDRWGQEYISGNEDPEGVWLKDLGPPMVARLVRFYPRADRVMSVCLRV	180	
Db	121	RYSRDGRRWGMKDRWGQEYISGNEDPEGVWLKDLGPPMVARLVRFYPRADRVMSVCLRV	180	
Qy	181	ELYGCLWRDGLLSYTPAVGQTYLSEAVYLNDSYDGTGVGGLQGLADGVVGLDD	240	
Db	181	ELYGCLWRDGLLSYTPAVGQTYLSEAVYLNDSYDGTGVGGLQGLADGVVGLDD	240	
Qy	241	FRSQEQLRVWPGYDYVGVSNHSSFSYGVMEFEEDRLRAFOAMOVHCNNMHTLGARLPGG	300	
Db	241	FRSQEQLRVWPGYDYVGVSNHSSFSYGVMEFEEDRLRAFOAMOVHCNNMHTLGARLPGG	300	
Qy	301	VECFRRGPGAMAVEGPMRHLNGNLGDPARAVSVPLGGRVARFLQCRFLFAGPWLIFS	360	
Db	301	VECFRRGPGAMAVEGPMRHLNGNLGDPARAVSVPLGGRVARFLQCRFLFAGPWLIFS	360	
Qy	361	EISFISDVWNSSPALGGTTPPAPWPPGPPPTNFSSLELEPRGQOPVAKPGESPTAILI	420	
Db	361	EISFISDVWNSSPALGGTTPPAPWPPGPPPTNFSSLELEPRGQOPVAKPGESPTAILI	420	
Qy	421	GCLVAITLLLLLIALMLRWRLLSKAERVLSEELTVHLSVPGDPTILNNRPGPRE	480	
Db	421	GCLVAITLLLLLIALMLRWRLLSKAERVLSEELTVHLSVPGDPTILNNRPGPRE	480	
Qy	481	PPYQEPERPRGNPPHAPCPVNGSALLSNPARYLLLATYARPPRGPGPPTPAWAKPTNT	540	
Db	481	PPYQEPERPRGNPPHAPCPVNGSALLSNPARYLLLATYARPPRGPGPPTPAWAKPTNT	540	
Qy	541	QAYSGDYMEPEKGPAPLLPPPPQNSVPHYAEADIVTLQGVGTGNTYAVPALPPGAVGDGP	600	
Db	541	QAYSGDYMEPEKGPAPLLPPPPQNSVPHYAEADIVTLQGVGTGNTYAVPALPPGAVGDGP	600	
Qy	601	PRVDFPRSRURFEKLEGEGQFGEVHLCVEVSDPSDLYSLDFPLNVRKXGHPLLVAVKILRPD	660	
Db	601	PRVDFPRSRURFEKLEGEGQFGEVHLCVEVSDPSDLYSLDFPLNVRKXGHPLLVAVKILRPD	660	
Qy	661	ATKNA-----RNDFLKEVKIMSLKDPNIIRLLGVCVDDPLCMITDYWENGDLNQELS	714	
Db	661	ATKNA-SFLSRNDFLKEVKIMSLKDPNIIRLLGVCVDDPLCMITDYWENGDLNQELS	720	
Qy	715	AHQLEDKAAAGADGQQAAGPTISYPMLLHVAAQAISGMRYLATUNFVHRDLATRNCLV	774	
Db	721	AHQLEDKAAAGADGQQAAGPTISYPMLLHVAAQAISGMRYLATUNFVHRDLATRNCLV	780	
Qy	775	GENETIKIADFGMSRNLVAGDYRVQGRVLPIRWMAWECILMGKFTTASDYVAFGVTLW	834	
Db	781	GENETIKIADFGMSRNLVAGDYRVQGRVLPIRWMAWECILMGKFTTASDYVAFGVTLW	840	
Qy	835	EVLMLCRAQFPGLTDEQVIENAGEFFRDQGRQVYLSRRPACPGQGLYELMLRCWSRESEQ	894	
Db	841	EVLMLCRAQFPGLTDEQVIENAGEFFRDQGRQVYLSRRPACPGQGLYELMLRCWSRESEQ	900	
Qy	895	RPFSQHLRFLAEDALNTV	913	
Db	901	RPFSQHLRFLAEDALNTV	919	

## RESULT 4

R71100  
ID R71100 standard; Protein; 914 AA.

AC	R71100;
AD	L7-AUG-1995 (first entry)
DE	Protein-tyrosine-kinase PTK22.
KW	Protein-tyrosine-kinase; PTK; discoidin domain receptor; cancer;
KW	breast tumor; mamma carcinoma; diagnosis; prognosis; therapy.
OS	Homo sapiens.
PN	WO9502187-A.
PD	19-JAN-1995.
PF	08-JUL-1994: G01480.
PR	09-JUL-1993: GB-014271.
PA	(CANC-) CANCER RES INST.
PA	(WELL) WELLCOME FOUND LTD.
PI	Barker KT, Crompton MR, Gusterson BA, Martindale JE;
PI	Mitchell PJ, Page MJ, Spence P;
DR	WPI: 95-066991/09.
DR	N-PSDB; Q84782.
PT	Method for screening substances, using protein tyrosine kinase -
PT	for potential utility as therapeutic agents for cancer
PS	Disclosure; Page 26-30; 51pp; English.
CC	CNA derived from tumor metastatic tissue was amplified using
CC	primers (given in Q84783-84) based on sequences (R71101, R71103)
CC	associated with protein-tyrosine-kinases (PTK). Novel PTK22 was
CC	identified in an isolated subclone. The 3' sequence of Q84786 a
CC	obtained by reverse transcription (using the primer of Q84786) a
CC	PCR amplification (primers Q84787-88) of RNA of human breast
CC	carcinoma cell line MDA MG 468. The partial DNA sequence of PTK
CC	is given in Q84782.
CC	Sequence 914 AA;
SQ	

Query Match 99.0%; Score 4852.5; DB 1; Length 914;  
Best Local Similarity 99.2%; Pred. No. 0;

Qy	1	MGPEALSSLLLLLVASGDADMGKHPDPAKCRVALGMQDRTIPDSDISASSWSNSTAAR	60
Db	1	MGPEALSSLLLLLVASGDADMGKHPDPAKCRVALGMQDRTIPDSDISASSWSNSTAAR	60
Qy	61	HSRLESSDGDGAWCPAGSVFPKKEEYLQVDLQRLHLVALVGTQGRHAGGLGKEFSSRYL	120
Db	61	HSRLESSDGDGAWCPAGSVFPKKEEYLQVDLQRLHLVALVGTQGRHAGGLGKEFSSRYL	120
Qy	121	RSRGRRWGKDRWQOEYISGNDEPEGVYLKDLGPPMVARLVREYPRADRVMVSYCLR	180
Db	121	RSRGRRWGKDRWQOEYISGNDEPEGVYLKDLGPPMVARLVREYPRADRVMVSYCLR	180
Qy	181	ELYCLWRDGLLSYTPAVPGQTMYLSEAVYLNDSITYDGHVTGGLOYGGLGQADGVVGLDD	240
Db	181	ELYCLWRDGLLSYTPAVPGQTMYLSEAVYLNDSITYDGHVTGGLOYGGLGQADGVVGLDD	240
Qy	241	FRKSOELRVMPGYDYVGSWNHSSFGVMEFEEDRLRFAQAMQVHCNNMHTLGARLPGG	300
Db	241	FRKSOELRVMPGYDYVGSWNHSSFGVMEFEEDRLRFAQAMQVHCNNMHTLGARLPGG	300
Qy	301	VECFRFRGPAMAWEGEPWRHNLGNLGDPRARAVSVPLGGRVAREFQCRLFLFAGPWLLFS	360
Db	301	VECFRFRGPAMAWEGEPWRHNLGNLGDPRARAVSVPLGGRVAREFQCRLFLFAGPWLLFS	360
Qy	361	EISFISDVMVNNSSPALGGTTPPAPWPPGPPPTNFSSLELEPRGQ-OPVAKPGESSTAIL	419
Db	361	EISFISDVMVNNSSPALGGTTPPAPWPPGPPPTNFSSLELEPRGQOPVAKAEGSSTAIL	420
Qy	420	IGCLVAITLLLLLLIALMLRHLHWRLLSKAERRVLEEELTVHLSVPGDPTILINNRPGR	479
Db	421	IGCLVAITLLLLLLIALMLRHLHWRLLSKAERRVLEEELTVHLSVPGDPTILINNRPGR	480
Qy	480	EPPTQEPRPGRNPPHSAPCVPNGSALLNPNAYRLLLATYARPPRGPGGPTTAWAKPTN	539
Db	481	EPPTQEPRPGRNPPHSAPCVPNGSALLNPNAYRLLLATYARPPRGPGGPTTAWAKPTN	540
Qy	540	TQAYSGDYMPEKFGAPLLPPPPONSVPHYAEADIVTLOGVTGGNTYAVPALPGVAGDG	599

```

Db 541 TQAYSGDYMEPRGAPLPPPPQNSVPHYAEADIVTLQGVGGNTYAYPALPPGAVGDG 600
QY 600 PPRVDFPRSLRKEKLGEGQFGEVHLCVEYDSDQDLVSLDFPLNVRKKGHPHLLVAVKILRP 659
Db 601 PPRVDFPRSLRKEKLGEGQFGEVHLCVEYDSDQDLVSLDFPLNVRKKGHPHLLVAVKILRP 660
QY 660 DATKNARNDLFKEVKTMSRLKDPNIIIRLIGVCVQDDPLCMITDYMENGLDQFLSAHQLE 719
Db 661 DATKNARNDLFKEVKTMSRLKDPNIIIRLIGVCVQDDPLCMITDYMENGLDQFLSAHQLE 720
QY 720 DKAAGAPGDGAQAQPTTSYPMLLHVAQAISGMRYLATLNFVHRDLATRNCLYGENFT 779
Db 721 DKAAGAPGDGAQAQPTTSYPMLLHVAQAISGMRYLATLNFVHRDLATRNCLYGENFT 780
QY 780 IKIADFGMSRLNLYAGDYRYVQGRVAVLPIRMAWECILMGKFTTASDVWAFGVTLWEVLM 839
Db 781 IKIADFGMSRLNLYAGDYRYVQGRVAVLPIRMAWECILMGKFTTASDVWAFGVTLWEVLM 840
QY 840 CQAQPFQGLTDEQVIENAGEFFRDQGRQVYLSRPPACPOGLYELMLRCWSRESEORPPFS 899
Db 841 CQAQPFQGLTDEQVIENAGEFFRDQGRQVYLSRPPACPOGLYELMLRCWSRESEORPPFS 900
QY 900 QLHRFLAEDALNTV 913
Db 901 QLHRFLAEDALNTV 914

RESULT 5
ID W34675
AC W34675;
DT 17-FEB-1998 (first entry)
DE Human mammary carcinoma kinase 10 (MCK-10) splice variant 3.
KW Mammary carcinoma kinase; MCK-10; receptor tyrosine kinase;
KW proliferative disease; cancer; insulin receptor family;
KW tyrosine kinase neurotrophin receptor; MCK-10 activity;
KW neurological disorder; aberrant expression.
OS Homo sapiens.
FH Key
FT 1. .18 Location/Qualifiers
FT Peptide
FT /label= signal_sequence
FT 19. .876
FT /note= "mature protein"
FT 31. .185
FT /label= Discoidin_I_like_domain
FT 304. .307
FT /label= endopeptidase_furin
FT /note= "putative precursor cleavage site"
FT 48. .439
FT /label= transmembrane_region
FT 580. .590
FT /label= ATP_binding_motif
FT 760. .761
FT /label= autophosphorylation_sites
FT /note= "putative"
FT 756. .756
FT /label= autophosphorylation_site
FT /note= "putative"
FT 802. .805
FT /label= binding_motif_for_pl3_kinase
FT /note= "binding motif for phosphatidylinositol 3'
FT 790
FT /label= potential_substrate_binding_site
FT 26. .42
FT /note= "antibody recognition sequence N1alpha"
FT 309. .321
FT /note= "antibody recognition sequence N1beta"
FT 860. .877
FT /note= "antibody recognition sequence CTbeta"
FT US5677144-A.
FT 14-OCT-1997.

```

```

PF 08-NOV-1994; 336343.
PR 16-NOV-1993; US-153397.
PA (ALVE/) ALVES F H E.
PA (ULLR/) ULLRICH A.
PI Alves FHE, Ullrich A;
DR WPI; 97-511869/47.
PT Truncated receptor tyrosine kinase CKK-2 - and nucleic acid coding
PS for it, useful for cancer diagnosis
PS Disclosure; Page 7; 70pp; English.
CC The present sequence represents a splice variant of a mammary
CC carcinoma kinase (MCK-10). This kinase belongs to a novel family
CC of receptor tyrosine kinases, and expression is associated with
CC proliferative diseases such as cancer. The MCK-10 receptor tyrosine
CC kinase has extensive sequence similarity to the insulin receptor family.
CC The MCK-10 gene was isolated by PCR using 2 degenerate oligonucleotide
CC primer pools, using a template cDNA synthesised by reverse transcription
CC of poly-A RNA from the human mammary carcinoma cell line MCF7. The
CC amplified PCR product was used to screen human foetal brain and placental
CC libraries, from which the present splice variant was isolated. This
CC splice variant does not possess amino acids 505-541 or 666-671 of MCK-10
CC (W34672). The sequence represented by amino acids 548-558 may be
CC important, as deletion of this motif in the activin receptor
CC serine/threonine kinase results in reduced ligand binding affinity.
CC MCK-10 is expressed in brain tissue, and the protein shares homology with
CC the tyrosine kinase neurotrophin receptor. Modulation of MCK-10 activity
CC therefore may be used for treatment of neurological disorders. MCK-10 is
CC also expressed in a variety of cancer cell lines and tumour tissue. The
CC nucleotide sequence of MCK-10, or parts of it, can be used for diagnostic
CC purposes to detect aberrant expression of MCK-10 genes. Inhibitors of
CC MCK-10 (or splice variants) receptor activity may have therapeutic value
CC in the treatment of diseases such as cancer.
CC note: the present sequence does not appear in the specification, but was
CC created using information provided.
SQ Sequence 876 AA;

```

Query Match 95.1%; Score 4664.5; DB 1; Length 876;

Best Local Similarity 95.8%; Pred. No. 0;

Matches 875; Conservative 0; Mismatches 1; Indels 37; Gaps 1;

```

QY 1 MGPEALSSLLLLLVASGDADMGHDPKACRYALGNQDRTIPDSISASSSSDSTAAR 60
DB 1 MGPEALSSLLLLLVASGDADMGHDPKACRYALGNQDRTIPDSISASSSSDSTAAR 60
QY 61 HSRLESSDGDGAWCPAGSVFPKEEYLVQDLQRLHLVALVGTQGRHAGGLGKFFSRYRL 120
DB 61 HSRLESSDGDGAWCPAGSVFPKEEYLVQDLQRLHLVALVGTQGRHAGGLGKFFSRYRL 120
QY 121 RYSDGRNMGKDRWGQEVISGNEDEPGVYLKDLGPPMVAVRLVFPYPRADRYMSVCLRV 180
DB 121 RYSDGRNMGKDRWGQEVISGNEDEPGVYLKDLGPPMVAVRLVFPYPRADRYMSVCLRV 180
QY 181 ELYGLWRDGLLSYTPVGTQMYLSEAVYLNDSTYDGHVTGGLOYGGLGADGVVGLDD 240
DB 181 ELYGLWRDGLLSYTPVGTQMYLSEAVYLNDSTYDGHVTGGLOYGGLGADGVVGLDD 240
QY 241 FRKSQELRVWPGDYDYGVWSNHSFSSGYVNEFEFDRLARFAQMVHCNNHHTLCARLPGG 300
DB 241 FRKSQELRVWPGDYDYGVWSNHSFSSGYVNEFEFDRLARFAQMVHCNNHHTLCARLPGG 300
QY 301 VECFRFRGPAMAVEGEPMRHNLGNLGDPRARAVSVPLGGRVARFLQCRFLFAGPMLLFS 360
DB 301 VECFRFRGPAMAVEGEPMRHNLGNLGDPRARAVSVPLGGRVARFLQCRFLFAGPMLLFS 360
QY 361 EISFISDVVNSSPALGGTTPAPMPPPGPTNFSSELEPGQOPVAKPEGSPATILI 420
DB 361 EISFISDVVNSSPALGGTTPAPMPPPGPTNFSSELEPGQOPVAKPEGSPATILI 420
QY 421 GCLVAIILLLLIIIALMLWRLHWRRLLSKAERRVLEETLVHLSVPGDITLINRPGPRE 480
DB 421 GCLVAIILLLLIIIALMLWRLHWRRLLSKAERRVLEETLVHLSVPGDITLINRPGPRE 480
QY 481 PPYQEPFRPNPNSAPCPVNGSALLLSNPAYRLLLATYARPPRPGPPPTPAWAKPTNT 540

```

Db 481 PPPQEPNPNPHSPAPVNGS----- 505  
Qy 541 QAYSGDYMEPEKCAPLLPPPPQNSVPHYAEADIVTLQGVGGNTYAVPALPPGAVGDGP 600  
Db 505 -AYSGDYMEPEKCAPLLPPPPQNSVPHYAEADIVTLQGVGGNTYAVPALPPGAVGDGP 563  
Qy 601 PRVDFPRSLRFLKELGEGQGEVHLCEVDSPODLVSLDFPLNVRKGHPLLVAVKILRPD 660  
Db 564 PRVDFPRSLRFLKELGEGQGEVHLCEVDSPODLVSLDFPLNVRKGHPLLVAVKILRPD 623  
Qy 661 ATKARNDFLKEVKIMSLKDPNIIIRLLGVCVQDDPLCMITDMYENGDLNQLSAHQLED 720  
Db 624 ATKARNDFLKEVKIMSLKDPNIIIRLLGVCVQDDPLCMITDMYENGDLNQLSAHQLED 693  
Qy 721 KAAGCAPGDGQAAGPPTISYPMMLHVAQAQASGRYATLNFVHRDLATNCLVGENFTI 780  
Db 684 KAAGCAPGDGQAAGPPTISYPMMLHVAQAQASGRYATLNFVHRDLATNCLVGENFTI 743  
Qy 781 KIADFGMSRNLVYAGDYRVQGRAVLPTRWMAWECILMGKFTTASDVWAFGVTLWEVLMCL 840  
Db 744 KIADFGMSRNLVYAGDYRVQGRAVLPTRWMAWECILMGKFTTASDVWAFGVTLWEVLMCL 803  
Qy 841 RAQPFQGLTDEQVIENAGEPRDQGRQVYLSRPPACQGLYELMLRCWSRESQRPFSQ 900  
Db 804 RAQPFQGLTDEQVIENAGEPRDQGRQVYLSRPPACQGLYELMLRCWSRESQRPFSQ 863  
Qy 901 LHRFLAEDALNTV 913  
Db 864 LHRFLAEDALNTV 876  
RESULT 6  
W34673  
ID W34673 standard; Protein; 882 AA.  
AC W34673;  
DE 17-FEB-1998 (first entry)  
KW Mammary carcinoma kinase 10 (MCK-10) splice variant 1.  
KW Mammary carcinoma kinase; MCK-10; receptor tyrosine kinase;  
KW proliferative disease; cancer; insulin receptor family;  
KW tyrosine kinase neurotrophin receptor; MCK-10 activity;  
KW neurological disorder; aberrant expression.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT Peptide 1..18  
FT /label= signal\_sequence  
FT Protein 19..919  
FT /note= "mature\_protein"  
FT Domain 31..185  
FT /label= Discoidin\_I\_like\_domain  
FT Cleavage\_site 304..307  
FT /note= "putative precursor cleavage site"  
FT Region 48..439  
FT /label= transmembrane\_region  
FT Binding\_site 580..590  
FT /label= ATP\_binding\_motif  
FT Modified\_site 765..766  
FT /label= autophosphorylation\_sites  
FT /note= "putative"  
FT Modified\_site 761  
FT /label= autophosphorylation\_site  
FT /note= "putative"  
FT Binding\_site 807..810  
FT /label= binding\_motif\_for\_P13\_kinase  
FT /note= "binding motif for phosphatidylinositol 3'  
FT kinase"  
FT Binding\_site 795..795  
FT /label= potential\_substrate\_binding\_site  
FT Region 26..42  
FT /note= "antibody recognition sequence NTalpha"  
FT Region 309..321  
FT /note= "antibody recognition sequence NTbeta"

FT Region 865..882  
FT US677144-A.  
FT 14-OCT-1997.  
FT 08-NOV-1994; 336343  
FT 16-NOV-1993; US-153397.  
FT (ALVE/) ALVES F H E.  
FT (ULR/) ULLRICH A.  
FT PI Alves FHE, Ullrich A;  
FT WPI: 97-511869/47.  
FT Truncated receptor tyrosine kinase CCK-2 - and nucleic acid coding  
FT for it, useful for cancer diagnosis  
FT Disclosure: Page -: 70pp; English.  
FT The present sequence represents a splice variant of a mammary  
FT carcinoma kinase (MCK-10). This kinase belongs to a novel family  
FT of receptor tyrosine kinases, and expression is associated with  
FT proliferative diseases such as cancer. The MCK-10 receptor tyrosine  
FT kinase has extensive sequence similarity to the insulin receptor family.  
FT The MCK-10 gene was isolated by PCR using 2 degenerate oligonucleotide  
FT primer pools, using a template cDNA synthesised by reverse transcription  
FT of poly-A RNA from the human mammary carcinoma cell line MCF7. The  
FT amplified PCR product was used to screen human foetal brain and placental  
FT libraries, from which the present splice variant was isolated. This  
FT splice variant does not possess amino acids 505-541 of MCK-10 (W34672).  
FT The sequence represented by amino acids 548-558 may be important, as  
FT deletion of this motif in the activin receptor serine/threonine kinase  
FT results in reduced ligand binding affinity. MCK-10 is expressed in brain  
FT tissue, and the protein shares homology with the tyrosine kinase  
FT neurotrophin receptor. Modulation of MCK-10 activity therefore may be used  
FT for treatment of neurological disorders. MCK-10 is also expressed in a  
FT variety of cancer cell lines and tumour tissue. The nucleotide sequence  
FT of MCK-10, or parts of it, can be used for diagnostic purposes to detect  
FT aberrant expression of MCK-10 genes. Inhibitors of MCK-10 (or splice  
FT variants) receptor activity may have therapeutic value in the treatment  
FT of diseases such as cancer.  
FT CC note: the present sequence does not appear in the specification, but was  
FT created using information provided.  
FT Sequence 882 AA;

Query Match 94.9%; Score 4651.5; DB 1; Length 882;  
Best Local Similarity 95.2%; Pred. No. 0;  
Matches 875; Conservative 0; Mismatches 1; Indels 43; Gaps 2;  
QY 1 MGPEALSSLLLLLVASGDADMKGHFDPKARYALGMQDRTPDSDISASSWSNSTAAR 60  
Db 1 MGPEALSSLLLLLVASGDADMKGHFDPKARYALGMQDRTPDSDISASSWSNSTAAR 60  
QY 61 HSRLESSDGDGAWCPAGSVFPKKEEYLQVDLQRLHLVALVGTQGRHAGGLGKEFSRYRL 120  
Db 61 HSRLESSDGDGAWCPAGSVFPKKEEYLQVDLQRLHLVALVGTQGRHAGGLGKEFSRYRL 120  
QY 121 RYSRGRRWGKRWGQGVISGNEDPEGVYLKDLGPPWVARLVRFYPRADRVMSVCLRV 180  
Db 121 RYSRGRRWGKRWGQGVISGNEDPEGVYLKDLGPPWVARLVRFYPRADRVMSVCLRV 180  
QY 181 ELYGCLWRDGLLSYAPVGTQMYLSEAVYLNDSYDGTHTVGGYGLGQADGVVGLDD 240  
Db 181 ELYGCLWRDGLLSYAPVGTQMYLSEAVYLNDSYDGTHTVGGYGLGQADGVVGLDD 240  
QY 241 FRKSOELRVNPGYDYVGNHSHFSSGYVMEFEFDRLRAFMQVHCNMHTLGLARLPGG 300  
Db 241 FRKSOELRVNPGYDYVGNHSHFSSGYVMEFEFDRLRAFMQVHCNMHTLGLARLPGG 300  
QY 301 VECFRFRGPAWAGEPEMRHNLGNLGDPPRARAVSVPLGGRVAREFLQCRFLFAGPWLIFS 360  
Db 301 VECFRFRGPAWAGEPEMRHNLGNLGDPPRARAVSVPLGGRVAREFLQCRFLFAGPWLIFS 360  
QY 361 EISFISDVYNNSSPALGGTFPPAPWPPGPPPTNFSSLELEPRGQOPVAKPGSPTAILI 420  
Db 361 EISFISDVYNNSSPALGGTFPPAPWPPGPPPTNFSSLELEPRGQOPVAKPGSPTAILI 420  
QY 421 GCIVAILLLLLLIIALLMLRLHWRLLSKARRVLEELITVHLSVPGDTILINNRPGPRE 480

```

Db 421 GCLVAIIILLLLIIALMLRLHWRLLSKAERVLSEELTVHLSVPGDTILINNRPGRPE 480
QY 481 PPYPQPRPRGPNPPHAPCVNPGNSALLSNPAYRLLLATYAPRPGPGPTPAWAKPTNT 540
Db 481 PPYPQPRPRGPNPPHAPCVNPGNS-----505
QY 541 QAYSGDYMEPEKPGAPLLPPPPQNSVPHYAEADIVTLQGTGNTYAVPALPGVAGDGP 600
Db 505 -AYSGDYMEPEKPGAPLLPPPPQNSVPHYAEADIVTLQGTGNTYAVPALPGVAGDGP 563
QY 601 PRYDFRSRLRFEKLGEGQFGEVHLCVDSPODLVSLDFPLNVRKGHPHLLVAVKILRPD 660
Db 564 PRYDFRSRLRFEKLGEGQFGEVHLCVDSPODLVSLDFPLNVRKGHPHLLVAVKILRPD 623
QY 661 ATKNA-----RNDLFKEVKIMSRKDPNIIRLLGVCVODDPLCMITDYMENGDLNQFLS 714
Db 624 ATKNASFSLFRNDFLKEVKIMSRKDPNIIRLLGVCVODDPLCMITDYMENGDLNQFLS 583
QY 715 AHOLEDKAAGAPGDCQAAGQPTISYPMLLHVAQAQASGRYLATLNFVHRDLATRNCLV 774
Db 684 AHOLEDKAAGAPGDCQAAGQPTISYPMLLHVAQAQASGRYLATLNFVHRDLATRNCLV 743
QY 775 GENFTIKIADFGMSRLYAGDYRVGQRAVLPIRMMAWECILMGKFTTASDVWAFGVTLW 834
Db 744 GENFTIKIADFGMSRLYAGDYRVGQRAVLPIRMMAWECILMGKFTTASDVWAFGVTLW 803
QY 835 EVLMLCRAQFPQGLTDEQVIENAGFEFRDQGRQVYLSRPPACPGQIYELMLRCWSESEQ 894
Db 804 EVLMLCRAQFPQGLTDEQVIENAGFEFRDQGRQVYLSRPPACPGQIYELMLRCWSESEQ 863
QY 895 RPPFSQLHRFLAEDALNTV 913
Db 864 RPPFSQLHRFLAEDALNTV 882

RESULT 7
W34674
ID W34674 standard; Protein: 563 AA.
AC W34674;
DT 17-FEB-1998 (first entry)
DE Human mammary carcinoma kinase 10 (MCK-10) splice variant 2.
KW Mammary carcinoma kinase; MCK-10; receptor tyrosine kinase;
KW proliferative disease; cancer; insulin receptor family;
KW tyrosine kinase neurotrophin receptor; MCK-10 activity;
KW neurological disorder; aberrant expression.
OS Homo sapiens.
FH Key Location/Qualifiers
FT 1..18
FT Peptide /label= signal_sequence
FT Protein 19..919
FT /note= "mature_protein"
FT Domain 31..185
FT /label= Discoidin_I_like_domain
FT Cleavage_site 304..307
FT /label= endopeptidase_furin
FT /note= "putative precursor cleavage site"
FT Region 48..439
FT Binding_site 617..627
FT /label= ATP_binding_motif
FT Modified_site 797..798
FT /label= autophosphorylation_sites
FT /note= "putative"
FT Modified_site 793
FT /label= autophosphorylation_site
FT /note= "putative"
FT Binding_site 839..842
FT /label= binding_motif_for_p13_kinase
FT /note= "binding motif for phosphatidylinositol 3' kinase"
FT Binding_site 827..827
FT /label= potential_substrate_binding_site

```

```

FT Binding_site 506..509
FT /label= putative_receptor_binding_site_for_SHC
FT /note= "SHC is an oncogenic SH2 domain containing molecule"
FT Binding_site 510..513
FT /label= GTPase_activity_protein_binding_site
FT /note= "putative"
FT Region 26..42
FT /note= "antibody recognition sequence NTalpha"
FT Region 309..321
FT /note= "antibody recognition sequence NTbeta"
FT Region 897..913
FT /note= "antibody recognition sequence CTbeta"
PN US5677144-A.
PD 14-OCT-1997.
PF 08-NOV-1994; 336343.
PR 16-NOV-1993; US-153397.
PA (ALVE)/ ALVES F H E.
PA (ULLR)/ ULLRICH A.
PI Alves FHE, Ullrich A;
DR WPI; 97-511869/47.
PT Truncated receptor tyrosine kinase CCK-2 - and nucleic acid coding
PS for it, useful for cancer diagnosis
PS Disclosure; Page -: 70pp; English.
CC The present sequence represents a splice variant of a mammary
CC carcinoma kinase (MCK-10). This kinase belongs to a novel family
CC of receptor tyrosine kinases, and expression is associated with
CC proliferative diseases such as cancer. The MCK-10 receptor tyrosine
CC kinase has extensive sequence similarity to the insulin receptor family.
CC The MCK-10 gene was isolated by PCR using 2 degenerate oligonucleotide
CC primer pools, using a template cDNA synthesised by reverse transcription
CC of poly-A RNA from the human mammary carcinoma cell line MCF7. The
CC amplified PCR product was used to screen human foetal brain and placental
CC libraries, from which the present splice variant was isolated. This
CC splice variant does not possess amino acids 566-671 of MCK-10 (W34672).
CC The sequence represented by amino acids 585-595 may be important, as
CC deletion of this motif in the activin receptor serine/threonine kinase
CC results in reduced ligand binding affinity. MCK-10 is expressed in brain
CC tissue, and the protein shares homology with the tyrosine kinase
CC neurotrophin receptor. Modulation of MCK-10 activity therefore may be used
CC for treatment of neurological disorders. MCK-10 is also expressed in a
CC variety of cancer cell lines and tumour tissue. The nucleotide sequence
CC of MCK-10, or parts of it, can be used for diagnostic purposes to detect
CC aberrant expression of MCK-10 genes. Inhibitors of MCK-10 (or splice
CC variants) receptor activity may have therapeutic value in the treatment
CC of diseases such as cancer.
CC note: the present sequence does not appear in the specification, but was
CC created using information provided.
SQ Sequence 563 AA;

```

```

Query Match 61.4%; Score 3011; DB 1; Length 563;
Best Local Similarity 99.8%; Pred. No. 2e-231;
Matches 562; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 351 LFAGPMLLFSEISFISDVNNSSPAGLGGTFPPAPWPPGPPPTNFSSLEPRGQGPVAK 410
Db 1 LFAGPMLLFSEISFISDVNNSSPAGLGGTFPPAPWPPGPPPTNFSSLEPRGQGPVAK 60
QY 411 PEGSPTAILGCLVAIIILLIIALMLRLHWRLLSKAERVLSEELTVHLSVPGDTI 470
Db 61 AEGSPTAILGCLVAIIILLIIALMLRLHWRLLSKAERVLSEELTVHLSVPGDTI 120
QY 471 LINNRPGPREPPYQPRPRGNPPHAPCVNPGNSALLSNPAYRLLLATYAPRPGPGPP 530
Db 121 LINNRPGPREPPYQPRPRGNPPHAPCVNPGNSALLSNPAYRLLLATYAPRPGPGPP 180
QY 531 TPAAWKPTNTQAYSGDYMEPEKPGAPLLPPPPQNSVPHYAEADIVTLQGTGNTYAVPA 590
Db 181 TPAAWKPTNTQAYSGDYMEPEKPGAPLLPPPPQNSVPHYAEADIVTLQGTGNTYAVPA 240
QY 591 LPPGAVGDGPPRVDFRSLRFEKLGEGQFGEVHLCVDSPODLVSLDFPLNVRKGHP 650

```



PT derived vectors, transformed cells, proteins and antibodies, useful  
PT for diagnosis and treatment of proliferative and nervous system  
PT diseases and for screening modulators  
PS Disclosure: Page 74-77; 115pp; English.  
CC A member of the mammary carcinoma kinase 10 (MCK-10) receptor  
CC tyrosine kinase family was identified using a PCR (with two  
CC degenerate oligo primer pools based on conserved sequences of the  
CC kinase domains of receptor tyrosine kinases) and cDNA prep. from  
CC colonic adenocarcinoma RNA. The nt sequence of the novel receptor,  
CC designated CCK-2, is given in Q92523 and the deduced AA sequence in  
CC R75505. Analysis of CCK-2 nt and AA sequences indicated significant  
CC homology with MCK-10 throughout the extracellular, transmembrane  
CC and intracellular regions. The regions of homology extend into the  
CC N-terminus consensus sequence of the discoidin I like family of  
CC proteins. CCK-2 was predominantly found in all stromal cells  
CC whereas MCK-10 expression was strongly confined to neoplastic  
CC cells themselves. Between the two RTKs, the juxtamembrane region  
CC is the region of most extensive sequence divergence.  
SQ Sequence 855 AA;

Query Match 49.3%; Score 2415; DB 1; Length 855;  
Best Local Similarity 52.1%; Pred No. 8.8e-184;  
Matches 492; Conservative 118; Mismatches 227; Indels 98; Gaps 15;

QY 3- PEALSLLLLLVASGDADMKGHFPAKCRVYALGMQDRTIPDSISASSWSDSAAARHS 62  
DB 5 PRMLLVLELLPLTS---SAKAQVNPACRYPLGMSGGQIPDEDITASSOWSESTAAYG 61  
QY 63 RLESSGDGAWCPAGSVFPKE-EYLOVDLQRLHLVALVGTGCRHAGGLGKEFFSRSLR 121  
DB 62 RLDSEGDGAWCPAEPVDPDLKEFLQIDLHTLFTLLVGTGCRHAGGHGIEFAPMYKIN 121  
QY 122 YSRDGRWGWKDRWGQEVISGNEDEGVVLDLGPVAVRLVRYPRADRVMSCLARVE 181  
DB 122 YSRDGRWISWRNKGKQVLDGNSNYDIFLKDLEPIVAVRFVPTVDHNSNVMCRVE 181  
QY 182 LYGLWRDLGSLTAPVPGQTMVL--SEAVYLDNDYDHTVGLGQYGLGLQADGVGLD 239  
DB 182 LYGCWLDGLVSYNAPAGQOQVLPGGSIYLDNDYVDG-AVGYSMTGLGQLTGCVGLD 240  
QY 240 DFRKSOELRWPGYDYVGNHSNHSFSGYVMEFEFDRLRAFOAMQVHCNMHTLGAIRLP 299  
DB 241 DFTQTHEYHWMPGYDYVGNHNSATNGYIEIMFEDRIENFTMTKVNHCNMFPAKGVKIF 300  
QY 300 GVCRRFRGPMAWECEPMRNLGNLGPDRARAVSVPLGVRAREFLOCRFLFAGPMLLF 359  
DB 301 EVOCYF-RSEASEWEPNAISFFPLVDVNPASRFVTPVPLHHRWASAIRKQYHFAFTWMWF 359  
QY 360 SEISFISD-VYNNSSPALGGTTPPAPWPPGPPPTNFSSLELEPRGQOQVAKPEGSPTAI 418  
DB 360 SEITFOSDAAMYNNSEAL----PTSP-----MAPTYDPLKAVDSDNTRI 400  
QY 419 LICLIVAILLLLIIITALLMLRWLRLLSKARRVLEELVHLVSPGDTILNRR--P 476  
DB 401 LICLIVAILFILLIIVIIILVWQKLEKASRRMLDDEMTVLSLPSDSSMNNRRSS 460  
QY 477 GPREP-----PPYQEPFRGNPHPSAPCVNPGSALLLSNPAYRLLLATYARP 523  
DB 461 SPSEQSNSTYDRIPLRPDYQEP-----SLRIKLPEF----- 495  
QY 524 PRGPGPTTFAWKATNTQAYSGDMEPEKPGAPLLPPPPQNSVPHYAEADIVTLQGVGG 583  
DB 495 -----APGEESGSCGVKVPQVSGP-----EGVPHYAEADIVNLOGVGG 535  
QY 584 NTAVPALPGVAGDGPVRY-DPPRSLRLEKELGEGQFGEVHLCEVDSQDVLSDPL 642  
DB 536 NTYSVPAVTMDLLSKDVAVEEPRLKATKEKLGEGQFGEVHLCEVEGMEKFKDKDFAL 595  
QY 643 NVKRGHPLLVAVKILRPDATTNARNDFLKEVKTMSRLKDPNIRLLGVCVQDDPLCMITD 702  
DB 596 DVSANQVPLVAVKMLRADANKARNDFLKEIKIMSLKDPNIIHVLGVCITDDPLCMITE 655

QY 703 YMENGDLNQFLSAHQLEDKAAEGAPGDGQAAOQTTSYPMLLHVAQAQIAGSMRYLATLNF 762  
DB 656 YMENGDLNQFLSRHE-----PPNSSSSSDVRYTVSYTNLKFATQIASGMKYLSSLNF 706  
QY 763 VHRDLATRNCLVGENFTIKIADFGMSRNLYAGDIYRVOGRAVLPIRMAWECILMGKFTT 822  
DB 707 VHRDLATRNCLVGNKTYIKIADFGMSRNLYSGDYIYIOGRAVLPIRMSWESILLGKFTT 766  
QY 823 ASDVWAFGVTLWEVLMCLCRAOPFGQLTDEQVIENAGFEFFRDQGRQVYLSRPPACPOGLYE 882  
DB 767 ASDVWAFGVTLWEVTFFCQEPYSQLSDEQVIENAGFEFFRDQGRQVYLPQPAICPDSVYK 826  
QY 883 LMLRCWSESEQRPPFSQLHFLAE 907  
DB 827 LMLSCWRRTKRNPSQOEIHLULLQ 851  
RESULT 10  
W34671 ID W34671 standard; Protein; 855 AA.  
AC W34671;  
DT 16-FEB-1998 (first entry)  
DE CCK-2, a human mammary carcinoma kinase 10 MCK-10) family member..  
KW Mammary carcinoma kinase; MCK-10; CCK-2; receptor tyrosine kinase;  
KW Proliferative disease; cancer; MCK-10 activity; aberrant expression.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT Modified\_site 121 /note= "N-glycosylated"  
FT Modified\_site 213 /note= "N-glycosylated"  
FT Modified\_site 261 /note= "N-glycosylated"  
FT Modified\_site 280 /note= "N-glycosylated"  
FT Modified\_site 328 /note= "N-glycosylated"  
FT Modified\_site 372 /note= "N-glycosylated"  
FT Modified\_site 503 /note= "N-glycosylated"  
FT Modified\_site /note= "putative autophosphorlation and substrate binding site"  
FT Modified\_site 736 /note= "putative autophosphorlation and substrate binding site"  
FT Modified\_site 740 /note= "putative autophosphorlation and substrate binding site"  
FT Modified\_site 741 /note= "putative autophosphorlation and substrate binding site"  
FT Modified\_site 813 /note= "putative autophosphorlation and substrate binding site"  
FT Modified\_site 825 /note= "putative autophosphorlation and substrate binding site"  
FT Region 400..421 /label= transmembrane\_region  
FT Domain 30..185 /note= "putative"  
FT Binding\_site 433..438 /label= Discoidin\_I\_like\_domain  
FT /label= protein\_kinase\_C\_binding\_site  
FT /note= "putative"  
US5677144-A.  
14-OCT-1997.  
PD 08-NOV-1994; 336343.  
PR 16-NOV-1993; US-153397.  
PA (ALIVE/) ALVES F H E.  
PA (ULLR/) ULLRICH A.  
PI ALVES FHE, Ullrich A;  
WPI; 97-511869/47.



DR N-PSDB; T93784.  
PT truncated receptor tyrosine kinase CCK-2 - and nucleic acid coding  
PS for it, useful for cancer diagnosis  
PS Claim 5; Fig 3; 70pp; English.  
CC The present sequence represents the amino acid sequence of human CCK-2, a  
CC member of the mammary carcinoma kinase 10 (MCK-10, W34572) family of  
CC receptor tyrosine kinases. The protein contains a remarkably high  
CC number of proline residues arranged as PXXP or PXXP repeats, suggesting  
CC a random coil structure for the hydrophilic juxtamembrane region. This  
CC region is probably a major domain for interactions with cellular  
CC substrates and other regulatory proteins. Expression of CCK-2 is  
CC associated with proliferative diseases such as cancer. The CCK-2 gene was  
CC identified by PCR and a cDNA prepared from colonic adenocarcinoma RNA.  
CC CCK-2 is expressed in a wide variety of cancer cell lines and tumour  
CC tissue. The CCK-2 nucleic acids can be used for diagnostic purposes to  
CC detect aberrant expression of CCK-2 genes. Engineered cell lines,  
CC containing recombinant vectors with the present sequence, are useful for  
CC producing infectious retroviral particles. The cell lines may also be  
CC used to evaluate and screen drugs involved in CCK-2 activation and  
CC regulation.  
SQ Sequence 855 AA;

Query Match 49.3%; Score 2415; DB 1; Length 855;  
Best Local Similarity 52.1%; Pred. No. 8.8e-184;  
Matches 482; Conservative 118; Mismatches 227; Indels 98; Gaps 15;

Qy 3 PEALSSLLLLVAGSDADMGHDPKACRYALCMQDRTIPDSDISASSWSSTAAHRS 62  
Db 5 PRMLLVLLPILS---SAKAQVNPALCRYPGLMSGGQIPDEDITASSQWSESTAAKYG 61

Qy 63 RLESSDGDGAWCAGSVPFKE-EYLVQVDLQRLHLVALVGTQGRHAGGLGKFSRSYRLR 121  
Db 62 RLDSEGDGAWCPEIPVDPDLKEFLQIDLHTLFTLVGTQGRHAGGHGIEFAPWKIN 121

Qy 122 YSRDGRRMWGHKDRWGQEVISGNEDEPGVVLKDLGPPMVARLVRYFPRADRVMSVCLRVE 181  
Db 122 YSRDGRTRWISWRNRHGKQVLDGNSNPYDIFLKDLPEPVARFVRFTPTDHSNMVCMRVE 181

Qy 182 LYGLWRDGLLSYAPVGTQMYL--SEAVYLDNSTYDGHVVGLOGLADGVVGLD 239  
Db 182 LYGCWLDGLVSNAPAGQVFLPGGSIYLDNSVDYG-AVGYSMTEGLQGLTDGVSGLD 240

Qy 240 DFRKSQELRVWPGDYVYGVNSHFSYVMEFEEDRLRAFAQWQVHCNNMHTLGARLPG 299  
Db 241 DFTQTHEYHWPGDYVYGVWRNESATNGYIETMEFDIRFTNTKMKVHCNNMFAKGVKIEK 300

Qy 300 GVECRFRGPAMWEGEPMRHNLGSLGDPRAVSVPLGGRVAVFLQCFEAGFWLLF 359  
Db 301 EVQCYF-RSEASEWEPNAISFPLVLDVNPNSARFVTVPLHHRWASAICKQYHEADTWMMF 359

Qy 360 SEISFISD-VYNNSSPALGCTFPAPWPPGPPPTNPFSSLELEPRGQOPVAKPEGSPTAI 418  
Db 360 SEITFQSDAAMYNSEAL---PTSP-----MAPTYDPMKLVDDSNTRI 400

Qy 419 LIGCLVAIIILLIILALMWRHLWRLLSKAERVLVEELTVHLSVPGDTILINNR--P 476  
Db 401 LIGCLVAIFILLIIVILWRQFQWMLKASRMDDDEMTVSLSPSSMFFNNRNS 460

Qy 477 GPREP-----PYQEPRPNGPHSAPCPVNGSALLSNPAYRLLLIATYAP 523  
Db 461 SPSSQGSNSTYDRFPLRPDYQEP-----SLRKLPEP----- 495

Qy 524 PRGCPPTPAWAKPTNTQAYSGDYMEPEKPGAPLLPPPPQNSVPHYAEADIVTLQGVG 583  
Db 495 -----AFGEESGCSGVVQVQSGP-----EGVPHYAEADIVNLQGVG 535

Qy 584 NTVAVPALPGVAGDGPVRV-DPPRSRLREKEKLGEGQFGEVHLCVDSQDVLVSLDFPL 642  
Db 536 NTVSVPATVMDLLSGDKVAVEEFPRKLLTEKELGEGQFGEVHLCVEGMEKEFKDKDFAL 595

Qy 643 NVKRGHPLLVAVKTLRPDATKNARNDLFKEVKIMSLRKDKPNIIRLLGVCVQDDPLCMITD 702  
Db 643 NVKRGHPLLVAVKTLRPDATKNARNDLFKEVKIMSLRKDKPNIIRLLGVCVQDDPLCMITD 702

Db 596 DVSANQPVLVAVKMLRADANKNARNDFLKEIKIMSLRKDPNIIRLLSVICITDDPLCMITE 655

Qy 703 YMENGDLNQFLSAHOLEDKAAEGAPGQQAAGPTISYPMLLHVAAQIASGMRYLATLNF 762

Db 656 YMENGDLNQFLSRHE-----PPNSSSDVRTVSTNLFKFWATQIASGMKYLSSLNF 706

Qy 763 VHRDLATRNCLVGENFTIKIADFGMSRNLVYAGDYRYVQGRAVLPIRMWAMECIILMGKFTT 822

Db 707 VHRDLATRNCLVGNKTYTIKIDFGMSRNLVYAGDYRYVQGRAVLPIRMWAMECIILMGKFTT 766

Qy 823 ASDVAFWAGVTLWEVLMLCRAFPFQGLTDEQVLENAGEFFRDGGRQVILSRPPACQGLYE 882

Db 767 ASDVAFWAGVTLWEVTFQEQEYFQSLSDQVIENTGEFFRDGGRQVILSRPPACQGLYE 826

Qy 883 LMLRCWSRESQRPFPFSLHRLFAE 907

Db 827 LMLSCWRDRDTKRNPSFQIHLILLQ 851

RESULT 11  
W77114

ID W77114 standard; Protein; 855 AA.  
AC W77114;  
DT 16-NOV-1998 (first entry)  
DE Discoidin domain receptor 2 protein.  
KW Discoidin domain receptor; transformation; metastasis; collagen; ss;  
KW Cleidocranial dysplasia; Sickler syndrome; extracellular matrix; MMP-1.  
OS Homo sapiens.  
PN W09834954-A2.  
PD 13-AUG-1998.  
PF 05-FEB-1998; CA0093.  
PR 06-FEB-1997; US-041578.  
PA (MOUN) MOUNT SINAI HOSPITAL CORP.  
PI Pawson A, Vogel W;  
DR WPI: 98-447186/38.  
DR N-PSDB; V48292.  
PT Novel ligands of discoidin domain receptor tyrosine kinase,  
PT especially collagen - useful for treating e.g. metastasis,  
PT cleidocranial dysplasia or Sickler syndrome  
PS Disclosure: Fig 22a; 115pp; English.  
CC The discoidin domain receptor (DDR) can be used to identify and evaluate  
CC substances which affect DDR receptor tyrosine kinase signalling pathways  
CC in the cell. Compounds which modulate such signalling pathways can be  
CC used to alter transformation or metastasis in mammals, to treat  
CC conditions involving structural or functional deregulation of collagens,  
CC e.g. Cleidocranial dysplasia or Sickler syndrome, conditions requiring  
CC modulation of extracellular matrix synthesis, degradation or remodelling,  
CC or to treat conditions needing modulation of MMP-1 expression such as  
CC wound healing.  
SQ Sequence 855 AA;

Query Match 49.3%; Score 2415; DB 1; Length 855;  
Best Local Similarity 52.1%; Pred. No. 8.8e-184;  
Matches 482; Conservative 118; Mismatches 227; Indels 98; Gaps 15;

Qy 3 PEALSSLLLLVAGSDADMGHDPKACRYALCMQDRTIPDSDISASSWSSTAAHRS 62  
Db 5 PRMLLVLLPILS---SAKAQVNPALCRYPGLMSGGQIPDEDITASSQWSESTAAKYG 61

Qy 63 RLESSDGDGAWCAGSVPFKE-EYLVQVDLQRLHLVALVGTQGRHAGGLGKFSRSYRLR 121  
Db 62 RLDSEGDGAWCPEIPVDPDLKEFLQIDLHTLFTLVGTQGRHAGGHGIEFAPWKIN 121

Qy 122 YSRDGRRMWGHKDRWGQEVISGNEDEPGVVLKDLGPPMVARLVRYFPRADRVMSVCLRVE 181  
Db 122 YSRDGRTRWISWRNRHGKQVLDGNSNPYDIFLKDLPEPVARFVRFTPTDHSNMVCMRVE 181

Qy 182 LYGLWRDGLLSYAPVGTQMYL--SEAVYLDNSTYDGHVVGLOGLADGVVGLD 239  
Db 182 LYGCWLDGLVSNAPAGQVFLPGGSIYLDNSVDYG-AVGYSMTEGLQGLTDGVSGLD 240

Qy 240 DFRKSQELRVWPGDYVYGVNSHFSYVMEFEEDRLRAFAQWQVHCNNMHTLGARLPG 299  
Db 241 DFTQTHEYHWPGDYVYGVWRNESATNGYIETMEFDIRFTNTKMKVHCNNMFAKGVKIEK 300

Qy 300 GVECRFRGPAMWEGEPMRHNLGSLGDPRAVSVPLGGRVAVFLQCFEAGFWLLF 359  
Db 301 EVQCYF-RSEASEWEPNAISFPLVLDVNPNSARFVTVPLHHRWASAICKQYHEADTWMMF 359

Qy 360 SEISFISD-VYNNSSPALGCTFPAPWPPGPPPTNPFSSLELEPRGQOPVAKPEGSPTAI 418  
Db 360 SEITFQSDAAMYNSEAL---PTSP-----MAPTYDPMKLVDDSNTRI 400

Qy 419 LIGCLVAIIILLIILALMWRHLWRLLSKAERVLVEELTVHLSVPGDTILINNR--P 476  
Db 401 LIGCLVAIFILLIIVILWRQFQWMLKASRMDDDEMTVSLSPSSMFFNNRNS 460

Qy 477 GPREP-----PYQEPRPNGPHSAPCPVNGSALLSNPAYRLLLIATYAP 523  
Db 461 SPSSQGSNSTYDRFPLRPDYQEP-----SLRKLPEP----- 495

Qy 524 PRGCPPTPAWAKPTNTQAYSGDYMEPEKPGAPLLPPPPQNSVPHYAEADIVTLQGVG 583  
Db 495 -----AFGEESGCSGVVQVQSGP-----EGVPHYAEADIVNLQGVG 535

Qy 584 NTVAVPALPGVAGDGPVRV-DPPRSRLREKEKLGEGQFGEVHLCVDSQDVLVSLDFPL 642  
Db 536 NTVSVPATVMDLLSGDKVAVEEFPRKLLTEKELGEGQFGEVHLCVEGMEKEFKDKDFAL 595

Qy 643 NVKRGHPLLVAVKTLRPDATKNARNDLFKEVKIMSLRKDKPNIIRLLGVCVQDDPLCMITD 702  
Db 643 NVKRGHPLLVAVKTLRPDATKNARNDLFKEVKIMSLRKDKPNIIRLLGVCVQDDPLCMITD 702



Db 241 DTQTHEYVWPYGVYVGRNESATNGYIEIMFEFDRINFTTKVHCNNMFAGVKIFK 300  
QY 300 GVECFRRGPMAMWEGEPWRLNGLNGLDPRARAVSVPLGGRVARFLQCRFLFAGPWLIF 359  
Db 301 EVQCYF-RSEASEWEPNAISFPLVLDVNPASRVFVPLHHRMASAIKCOYHFADTWMF 359  
QY 360 SEISFISD-VYNNSSPALGGTFPPAPWPPGPPPTNFSSLEPRGQOPVAKPEGSPAI 418  
Db 360 SEITFQSDAAMYNNSEAL-PTSP-----MPTTYDPMKLVDDSNTRI 400  
QY 419 LGICLVAILILLILLIALLMLARLHWRLLSKAERVLLEELTVHLSVPGDTILNRR--P 476  
Db 401 LGICLVAILIFILLIIVILWRQVQWMLERASRMDDDEMVSLSLPSOSSMFNNRRSS 460  
QY 477 GPREP-----PPYQEPRPGRNPPHSPACVPNGSALLSNPRLLLATYARP 523  
Db 461 SPSEQGSNSTYDRIFPLRPDQEP-----SRLIRKLPEF----- 495  
QY 524 PRGPGPTPAWAKTNTQAYSGDYMEPEKPGAPLLPPPQNSVPHYAEADIVTLQGVGG 583  
Db 495 -----APGEEESGCGVVKVPQPSGP-----EGVPHYAEADIVNLQGVGG 535  
QY 584 NTYAVPAPPGAVGDPGRV-DEPRSLRREKELGEGQFGEVHLCVDSQDLVSLDFPL 642  
Db 536 NTYSPVATVMDLLSGKDVAVEEFPRKLLTFKEKLGEGQFGEVHLCVDSQDLVSLDFPL 595  
QY 643 NVKRGHPLLVAVKILRPDATKNARNDELKEVKIMSLRKDPNIIIRLLGVQDDPCLCMITD 702  
Db 596 DVSANOPVLVAVKMLRADANKARNDELKEIKIMSLRKDPNIIIRLLGVQDDPCLCMITE 655  
QY 703 YWENGDLNQFLSAHQLEDKAAEGAPGDGQAAGQPTISYPMLLHVAQAQIASGRYLAFLNF 762  
Db 656 YWENGDLNQFLSRHE-----PPNSSSDVTVSYTNLKFATQIASGMKYLSSNF 706  
QY 763 VHRDLATNCLVGENFTIKIADFGMSRLNLYAGDYRYVQGRAVLPIRMWAECLLMGKFTT 822  
Db 707 VHRDLATNCLVGNFTIKIADFGMSRLNLYSGDYRYVQGRAVLPIRMWAECLLMGKFTT 766  
QY 823 ASDVWAFVGLTWELVLMCRAPQFQGLTDEQVIENAGFEFFRQGVYLSRPPACPGLYE 882  
Db 767 ASDVWAFVGLTWELTFTFCQEQPYSQLSDEQVIENAGFEFFRQGVYLSRPPACPGLYE 826  
QY 883 LMLRCWSESEQRPPFSQLHRFLAE 907  
Db 827 LMLSCWRDRTKNRPSFQBIHLLILQ 851

## RESULT 12

W79152  
ID W79152 standard; Protein; 854 AA.  
AC W79152;  
DT 19-NOV-1998 (first entry)  
DE Receptor protein tyrosine kinase (PTK) subtype tyro-10.  
KW PTK; receptor; protein tyrosine kinase; brain tissue.  
OS Rattus sp.  
PN US581516-A.  
PD 22-SEP-1998.  
PF 15-MAY-1992; 456647.  
PR 02-JUN-1995;  
PR 15-MAY-1992; US-884486.  
PR 02-MAY-1994; US-237401.  
PR 02-JUN-1995; US-456647.  
PA (SALK ) SALK INST BIOLOGICAL STUDIES.  
PI Lai CHC, Lemke GE.  
DR WPI; 98-53039/45.  
DR N-PSDB; V55895.  
PT Receptor protein tyrosine kinase polypeptide, tyro-3 - preferably expressed in brain tissue  
PS Example 2; Columns 53-60; 46pp; English.  
CC This represents a novel receptor protein tyrosine kinase (PTK) polypeptide subtype tyro-10. The invention provides polynucleotide sequences encoding novel PTK polypeptide subtypes tyro-1 to tyro-13.  
CC The PTK subtypes are found expressed predominantly in the brain tissue.

SQ Sequence 854 AA;

Query Match 49.2%; Score 2413; DB 1; Length 854;  
Best Local Similarity 52.3%; Pred. No. 1.3e-183;  
Matches 481; Conservative 119; Mismatches 220; Indels 100; Gaps 15;

QY 9 LLLLLLVAGSDADMKGHFDPAKRYALGDMODRTIPDSDISASSSSSDSTAARHSLESD 68  
Db 10 VLLLLLLILGSA--KAQVNPACIRPLGMSGGHIPEEDITASSQWSESTAARIGLDEE 67  
QY 69 GGAWCPAGSVFPKE--EYLOVDLQRLHLVALVGTQGRHAGGLGKFEFSRYRLRYSDRG 127  
Db 68 GGAWCPETPVQDDLEKFLQIDLRLHFTILVGTQGRHAGGHTIEFAPMYKINYSRDS 127  
QY 128 RWMGKDKRWGOEIVSNGEDPEGVWLKDLGPPMVARLVREYPRADRVMSVCLRVLYGLW 187  
Db 128 RWISWMNRHGKQVLDGNSNPYDFVKLDLEPPIVAREFLIPVTDHSMNVCMREVELYGVW 187  
QY 188 RDGLLSYTAPOVGTMYL--SEAVYLNDSYDGTGTVGLQYGLGLQADGVGLDDFRKSQ 245  
Db 188 LDGLVSYNAPAGQQFVLPGSGSIILNDSYDGTG--AVGYSTEGLGQLTDGVSLDFTQTH 246  
QY 246 ELRWMPGYDYVGNWSHSSFGVYEMEFEDRLRFAQMOVHCNNMHTLGLARLPGGVECRF 305  
Db 247 EYHVMPPGYDYVGNWRNESATNGFIEMFEFDRIRNFTTKVHCNNMFAGVKIFKEVQCYF 306  
QY 306 RGPAMAWEGEPWRLNGLNGLDPRARAVSVPLGGRVARFLQCRFLFAGPWLIFSEISFI 365  
Db 307 -RSEASEWEPNAISFPLVLDVNPASRVFVPLHHRMASAIKCOYHFADTWMFSEITFQ 365  
QY 366 SD--VYNNSSPALGGTFPPAPWPPGPPPTNFSSLEPRGQOPVAKPEGSPAILIGCL 423  
Db 366 SDAAMYNSS-----GALPTSP-----MPTTYDPMKLVDDSNTRIIGCL 405  
QY 424 VAILLLLLLIILMLWRHLRRLSKAERVLLEELTVHLSVPGDTIILNRR----PGPR 479  
Db 406 VAIIFILLIIVILWRQVQWMLERASRMDDDEMVSLSLPSOSSMFNNRRSSPSEQ 465  
QY 480 EP-----PYQEPRPGRNPPHSPACVPNGSALLSNPRLLLATYARPGRPG 528  
Db 466 ESNSTYDRIFPLRPDQEP-----SRLIRKLPEF----- 495  
QY 529 PPTPAWAKPTNTQAYSGDYMEPEKPGAPLLPPPQNSVPHYAEADIVTLQGVGGTNTYAV 588  
Db 495 -----APGEEESGCGVVKVPQPSGP-----EGVPHYAEADIVNLQGVGGTNTYAV 540  
QY 589 PALPPGAVGDPGRV-DFPRSLRREKELGEGQFGEVHLCVDSQDLVSLDFPLNVRKG 647  
Db 541 PAVTMDLLSGKDVAVEEFPRKLLAFKEKLGEGQFGEVHLCVDSQDLVSLDFPLNVRKG 600  
QY 648 HPLLVAVKILRPDATKNARNDELKEVKIMSLRKDPNIIIRLLGVQDDPCLCMITDYENG 707  
Db 601 QPVLVAVKMLRADANKARNDELKEIKIMSLRKDPNIIIRLLAVCITEDPCLCMITEYENG 660  
QY 708 DLNQFLSAHQLEDKAAEGAPGDGQAAGQPTISYPMLLHVAQAQIASGRYLAFLNVRHDL 767  
Db 661 DLNQFLSRHEPLSSCSDA-----TVSYANLKFATQIASGMKYLSSLNFVHDL 710  
QY 768 ATRNCLVGENFTIKIADFGMSRLNLYAGDYRYVQGRAVLPIRMWAECLLMGKFTTASDVW 827  
Db 711 ATRNCLVGNKNTIKIADFGMSRLNLYSGDYRYVQGRAVLPIRMWAECLLMGKFTTASDVW 770  
QY 828 AFGVTLWELVLMCRAPQFQGLTDEQVIENAGFEFFRQGVYLSRPPACPGLYELMLRC 887  
Db 771 AFGVTLWELTFTFCQEQPYSQLSDEQVIENAGFEFFRQGVYLSRPPACPGLYELMLRC 830  
QY 888 WSRSEQRPPFSQLHRFLAE 907  
Db 831 WRRETKHRPSFQBIHLLILQ 850

RESULT 13

W81409  
 ID W81409 standard; Protein: 854 AA.  
 AC W81409.  
 DT 22-JAN-1999 (first entry)  
 DE Receptor protein tyrosine kinase (PTK) subtype tyro-10.  
 KW PTK; receptor; protein tyrosine kinase; recombinant; grafting;  
 KW diagnosis; tumour; skin transplant; connective tissue; tyro-10.  
 OS Rattus sp.  
 PN US5937448-A.  
 PD 17-NOV-1998.  
 PF 02-MAY-1994; 237401.  
 PR 15-MAY-1992; US-884486.  
 PR 02-MAY-1994; US-237401.  
 PA (SALK ) SALK INST BIOLOGICAL STUDIES.  
 PI Lai CHC. Lemke GE.  
 DR WPI: 99-023436/02.  
 DR N-PSDB; V65317.  
 PT Nucleic acids encoding protein tyrosine kinase subtypes - for  
 PT identification of new subtypes and treatment of diseases associated  
 PT with the kinase  
 PS Claim 10: Columns 53-58: 47pp: English.  
 CC This represents a receptor protein tyrosine kinase (PTK) subtype tyro-10.  
 CC The invention provides sequences V65308 to V65313, V65315, and V65317 to  
 CC V65319 that encode proteins having a tyrosine kinase domain and a tissue  
 CC expression pattern of a receptor PTK subtype selected from tyro-1,  
 CC tyro-2, tyro-3, tyro-4, tyro-5, tyro-6, tyro-8, tyro-10, tyro-11, and  
 CC tyro-12, respectively. The polynucleotides are useful for the detection  
 CC of tyrosine kinase domain sequences and detection of tissue expression  
 CC patterns of PTK subtypes. The cDNAs can also be injected into oocytes,  
 CC the protein expressed, and expression products screened for using  
 CC antibodies against tyrosine kinase epitopes. These subtype sequences can  
 CC be used for the design of oligonucleotides, for use in amplification  
 CC reactions to isolate other subtype sequences. These detection protocols  
 CC are used in the diagnosis of diseases associated with (receptor) PTKs.  
 CC Recombinant vectors expressing the subtypes can be used to treat related  
 CC diseases e.g. tumours, by introduction of the vectors into skin  
 CC transplants, then grafting these into the connective tissue of the  
 CC dermis, thus specifically targeting tumours as the proteins are released  
 CC from the matrix.  
 CC Sequence 854 AA;  
 SQ

Query Match 49.2%; Score 2413; DB 1; Length 854;  
 Best Local Similarity 52.3%; Pred. No. 1.3e-183;  
 Matches 481; Conservative 119; Mismatches 220; Indels 100; Gaps 15;

QY 9 LLLLLLVASGDADKGFDPKCRYALGMQDRTIPSDISASSWSDSSTAARSLRSSD 68  
 DB 10 VLLLLLLILGSA--KAQVNPACRYPLGMSGGHIPDEDITASSQWSESTAARYGRLDSEE 67  
 QY 69 GDGAWCPAGSVFPKE-EYELQVDLQRLHLVALVGTQGRHAGGLGKFFRSYRLRYSRDGR 127  
 DB 68 GDGAWCEIPVQDDLKEFLQIDLRLHTLTLVGTQGRHAGHGIEFAPMYKINTSRDGS 127  
 QY 128 RWGKWDKRGQVSGNDEGVLKDLGPPMVARLVRYPRADRVMSVCLRVLYGLCLW 187  
 DB 128 RWISWRNRHKGQVLGDSNENYDFELKDLPEPIVAREVRLPIVTHDSNMVCMRVLYGCVW 187  
 QY 188 RDGLLSTAPVGTQMYL--SEAYLVNDSTVDGHTVGLQYVGLQGLADGVGLDDPRKSO 245  
 DB 188 LDGLVSNAPAGQOVLPGGSGIYLLNDSTVDG--AVGYSMTGELGQUTDGVGLDDFTQT 246  
 QY 246 ELRWPCYDYGVGNSHFSFGYVEMFEFFDLRLAFAMQVHCNMHTLGLARLPGGVECRF 305  
 DB 247 EYHWPCYDYGVGRNESATNGFIEFEDRIIRNFTTMKVHCNMFAKGVKIFKEVQCYF 306  
 QY 306 RRGPAWAGEPMRHNGLGNDPRARAVSVPLGGVAREFLQCRFLFAGPWLFFSISFI 365  
 DB 307 -RSEASEWPTAVYFVLVDVNPASRFVTVPLHHRMASAIKQCYHFDATWMMFSEITQ 365  
 QY 366 SD--VVNNSPALGGTTPPAPWPPPGPTTFSSLEPRGQOPVAKPGSPATLIGCL 423  
 DB 366 SDAAMYNNS-----GALTPSP-----MAPTYDPMKLVDDSNTRILIGCL 405

QY 424 VAITLLLLIIALMLRLHWRLLSKAERVRLEELTVHLSVPGDTILNNR----PGPR 479  
 DB 406 VAIIFILLAIIVILWRQFQWKLEKASRMDDDEMTVLSLPSESMNNRNSPSEQ 465  
 QY 480 EP-----PPYOEPRPRGNPPHSAFCVPGVNGSALLSNPAYLLATYARPPRGG 528  
 DB 466 ESNSTYDRIFPLRPDYQEP-----SRLIKLPEF----- 495  
 QY 529 PPTPAWAKPTNTQAYSQDYMEPEKPGAPLPPPPQNSVPHYAEADIVTLOGVTGGNTYAV 588  
 DB 495 -----APGEEESGCGVVKPAQPNPG-----EGVPHYAEADIVNLOGVTGGNTYCV 540  
 QY 589 PALPPGAVGDGPPRV-DFPRSRRLRFKEKLGEGQFGEVHLCVEVDSPQDLVSLDPLNVRKG 647  
 DB 541 PAYTMDLLSGKDVAVEEFPRKLLAFKEKLGEGQFGEVHLCVEGMEKFKDKDFALDVSAN 600  
 QY 648 HPLVAVKILRPDATKNARNDFLKEVKIMSRKDPNIIRLLGVCVQDDPLCMITDYMENG 707  
 DB 601 QPVLVAVKMLRADANKNARNDFLKEIKIMSRKDPNIIRLLAVCIITDPLCMITEYMEG 660  
 QY 708 DLNQFLSAHQLEDKAAEGAPGDCQAAGPTISYPMLLHVAQAQIASGMRYLATINEVHRDL 767  
 DB 661 DLNQFLSRHHEPLSCSSDA-----TVSYANLKFWATQASGMKYLUSSLNFEVHRDL 710  
 QY 768 ATRNCLVGENFTIKIADFGRNRLYAGDYRVQGRAVLPIRMMAWECILMGKFTTASDVW 827  
 DB 711 ATRNCLVGNKYTIKIADFGRNRLYAGDYRVQGRAVLPIRMMSWESILLGKFTTASDVW 770  
 QY 828 AFGVTLMVLMCRAQPGQLTDEQVITENAGFEFRQGRQVYLSRPPACQGYELMLRC 887  
 DB 771 AFGVTLMETTFCEQYPSQLSDEQVIENTGEFFRQGRQIYLPQALCPDSPYKMLMSC 830  
 QY 888 WRESEQRPPFSQLHRELA 907  
 DB 831 WRRETKHRPSFQEIHLHLQ 850

RESULT 14  
 R54089  
 ID R54089 standard; Protein: 650 AA.  
 AC R54089;  
 DT 03-FEB-1995 (first entry)  
 DE Partial sequence of tyrosine kinase receptor protein.  
 KW Tyrosine kinase; receptor; proto-oncogene; trk; detection;  
 KW diagnosis; antibody; treatment; tumour; antisense.  
 OS Homo sapiens.  
 FH key. Location/Qualifiers  
 FT modified\_site 8  
 FT modified\_site 56 /note= "Potential N-linked glycosylation site."  
 FT modified\_site 75 /note= "Potential N-linked glycosylation site."  
 FT modified\_site 167 /note= "Potential N-linked glycosylation site."  
 FT domain 195..216 /note= "Potential N-linked glycosylation site."  
 FT region 365..370 /label= Transmembrane domain.  
 FT /label= ATP binding region.  
 FT /note= "Highly conserved among protein tyrosine kinase enzymes."  
 FN DE4239817-A.  
 PD 01-JUN-1994.  
 PF 26-NOV-1992; 239817.  
 PR 26-NOV-1992; DE-239817.  
 PA (CHEM-) CHEMOTHERAPEUTISCHES FORSCHUNG.  
 PI Holtrich U, Ruebsamen-waigmann H, Streibhardt K;  
 DR WPI: 94-184380/23.  
 DR N-PSDB; 064158.  
 PT New protein tyrosine kinase and related nucleic acid - vectors,  
 PT transformed cells, etc., useful for diagnosis and treatment of  
 PT tumours

FT		/note="potential N-linked glycosylation site"
FT	misc_difference 241..244	/note="potential N-linked glycosylation site"
FT		/note="potential N-linked glycosylation site"
FT	misc difference 254..257	

FT	misc_difference	280. .283	/note= "potential N-linked glycosylation site"
FT			
FT	misc_difference	325. .328	/note= "potential N-linked glycosylation site"
FT			

```

FFI /note= "potential N-linked glycosylation site
FTT misc_difference 338. .341

```

```

/Note= "potential N-linked glycosylation site
FT      misc_difference 412. .415

```

FT domain 431..454

```

FI      Transmembrane domain
FI      /note= "Transmembrane domain"
FT      544..807
FT      domain

```

```

FI /note= "tyrosine-kinase domain"
FT misc_difference 466

```

EF1	/note= splice site for truncated ttrc"
PN	W09525795-A1.

FD 28-SEP-1995.  
PF 17-MAR-1995; U03426.

FR 18-MAR-1994; US-215139.  
PR 05-AUG-1994; US-286846.

PA (GETH ) GENENTECH INC.  
20 DEC-1994, US-339703.

FI Fiesta LG, Shellcon DE, Utter K;  
DR WPI; 95-344616/44.

PT New human trkB and trkC poly:peptide(s) and fusion proteins co-

PT and diagnosis of abnormal neurotrophic factor expression, e.g.

PS Claim 8; Fig.1A-B; 117pp; English.

CC of human *trkC* receptor, and to detect or amplify *trkC* genes. This sequence may be expressed recombinantly for the production of human *trkC* receptor, and to detect or amplify *trkC* genes.

CC activation assays, and therapeutically in diseases associated

over or under expression of neurotrophic factor (e.g. pain or inflammation, kidney, lung, cardiovascular or psychiatric disorders)

CC and some sorts of tumours). At the indicated splice site, AA  
CC 791 may be replaced by the sequence FVLFHKIPLDG' (R81624).

sequence  
022 nm,

Query Match 13.4%; Score 658; DB 1; Length 822;

Best local similarity 47.4%; Error: NO. 4.2E-44;  
Matches 144; Conservative 45; Mismatches 79; Indels 36;

QY 607 RSRLRFKEKLGEGQFGEVHLCE--VDSPQDLVSLDFPLNVRKGHPLLVAVKILRPD

Dbb 535 RHNIVLKRELGEAGFKVFLAECYNLCPEQD-----KILVAVKTLK-D

664 QY NARDFLKEVKIMSRKDPNIIRLLGVCVQDDPLCMITDYMENGDLNQFLSAHQLED

Dbb 580 NARKDFHREAE LLTNLQHEHIVFYGVCEVDPLIMVFYEMKHGDLNKF LRAHGPDA

QY 723 AEGAPGDQAAQGPT- ISYPMLLHVAAQIASGMRYLATLNFVHRDLATRNCVLVGENF

Dbb 640 AEGNP-----PTELTQSQMLHIAQQIAAGMVYLASQHFVHRDLATRNCLVGNL

QY 782 IADFGMSRNLYAGDYRVQGRAVLPIRMWAWECILMGKFTTASDVWAFGVTLWEVLM

Db 692 IGDFGMSRDVYSTDYRVGGHTMLPIRWMPESIMYRKFTTESDWSLGVWLWEIFT

Qy 842 AOPFGQLTDEQVIENAGEFFRDQGRQVYLSRPPACPOGLYELMLRCWSRESEQRPPFSQL 901  
Db 752 -QPWYQLSNNEVIECI-----TQGR--VLQRPRTCPQEVYELMLGCVQREPHMRKNIKGI 803  
Qy 902 HRFL 905  
Db 804 HTLL 807

Search completed: November 3, 1999, 22:52:09  
Job time: 3331 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run On: November 4, 1999, 07:05:18 ; Search time 752.95 Seconds  
(without alignments)  
3135.829 Million cell updates/sec

Title: US-08-170-558-7

Perfect score: 1197

Sequence: 1 GATGCTGACATCAGGGACA.....CCGAGGGAGCCCGACCGCC 1197

Scoring table: IDENTITY\_NUC

Searched: 2546578 seqs, 986266752 residues

Database :

EST:\*

1: em\_est1:.\*  
2: em\_est2:.\*  
3: em\_est3:.\*  
4: em\_est4:.\*  
5: em\_est5:.\*  
6: em\_est6:.\*  
7: em\_est7:.\*  
8: em\_est8:.\*  
9: em\_est9:.\*  
10: em\_est10:.\*  
11: em\_est11:.\*  
12: em\_est12:.\*  
13: em\_est13:.\*  
14: em\_est14:.\*  
15: em\_est15:.\*  
16: em\_est16:.\*  
17: em\_est17:.\*  
18: em\_est18:.\*  
19: em\_est19:.\*  
20: gb\_est1:.\*  
21: gb\_est2:.\*  
22: gb\_est3:.\*  
23: gb\_est4:.\*  
24: gb\_est5:.\*  
25: gb\_est6:.\*  
26: gb\_est7:.\*  
27: gb\_est8:.\*  
28: gb\_est9:.\*  
29: gb\_est10:.\*  
30: gb\_est11:.\*  
31: gb\_est12:.\*  
32: gb\_est13:.\*  
33: gb\_est14:.\*  
34: gb\_est15:.\*  
35: gb\_est16:.\*  
36: gb\_est17:.\*  
37: gb\_est18:.\*  
38: gb\_est19:.\*  
39: gb\_est20:.\*  
40: gb\_est21:.\*  
41: gb\_est22:.\*  
42: gb\_est23:.\*  
43: gb\_est24:.\*  
44: gb\_est25:.\*  
45: gb\_est26:.\*  
46: gb\_est27:.\*  
47: gb\_est28:.\*  
48: gb\_est29:.\*  
49: gb\_est30:.\*  
50: gb\_est31:.\*  
51: gb\_est32:.\*  
52: em\_est20:.\*  
53: em\_est21:.\*

54: em\_est22:.\*  
55: em\_est23:.\*  
56: em\_est24:.\*  
57: em\_est25:.\*  
58: em\_est26:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	474.4	39.6	772	54	HSM008101	AL043251 Homo sapi
2	314.8	26.3	433	28	AA073943	AA073943 mm97hl2.r
C 3	210.8	17.6	568	36	AA622035	AA622035 nq23b08.s
4	177.8	14.9	257	54	HSM011191	AL046341 Homo sapi
C 5	119.4	10.0	245	44	AI300297	AI300297 gm69hl2.x
6	116.8	9.8	235	23	R85358	R85358 yo39hl2.r1
7	101.6	8.5	316	20	T28305	T28305 EST37127.Hu
8	73.2	6.1	82	23	H38476	H38476 YP69C07.r1
9	46.2	3.9	495	37	AA734307	AA734307 vt26f03.r
10	45.8	3.8	515	49	AV000273	AV000273 AV000273
11	44.6	3.7	520	30	AA221816	AA221816 my28f05.r
12	44.6	3.7	503	35	AA597143	AA597143 vo35f03.r
13	44.6	3.7	484	37	AA691209	AA691209 vt34e06.r
14	44.6	3.7	507	37	AA691386	AA691386 vu15a10.r
15	44.6	3.7	518	37	AA692234	AA692234 vt23h01.r
16	44.6	3.7	596	39	AA815999	AA815999 vr14b12.r
17	44.2	3.7	335	38	AA781615	AA781615 ai56f01.s
C 18	44.2	3.7	530	39	AA857320	AA857320 oh97a11.s
C 19	44.2	3.7	441	40	AA991406	AA991406 os52d12.s
20	44.2	3.7	243	49	AU058693	AU058693 AU058693
21	44.2	3.7	243	49	AU058845	AU058845 AU058845
22	44.2	3.7	243	49	AU058851	AU058851 AU058851
C 23	44	3.7	501	36	C77167	C77167 C77167 Mous
24	43.8	3.7	459	23	D52589	D52589 HUM082810B
25	43.8	3.7	243	49	AU059967	AU059967 AU059967
C 26	43.2	3.6	453	44	AA289917	AA289917 qv02el2.x
C 27	43	3.6	427	28	AA084238	AA084238 zn04h06.s
C 28	43	3.6	565	28	AA108143	AA108143 ml96b10.r
C 29	43	3.6	489	34	AA499317	AA499317 vi92h05.r
C 30	43	3.6	500	37	AA734118	AA734118 vv22f10.r
C 31	43	3.6	736	39	AA822411	AA822411 vw37d09.r
C 32	43	3.6	781	42	AI092944	AI092944 qa81c01.x
C 33	43	3.6	722	42	AI098196	AI098196 ue31c12.x
C 34	43	3.6	495	47	AI538451	AI538451 td06d12.x
C 35	43	3.6	456	49	AV001579	AV001579 AV001579
C 36	42.8	3.6	466	29	AA150564	AA150564 zo89a04.s
C 37	42.6	3.6	419	34	AA468397	AA468397 nc78g07.s
C 38	42.6	3.6	568	35	AA587115	AA587115 nn70a10.s
C 39	42.6	3.6	487	38	AA738127	AA738127 nx16e01.s
C 40	42.6	3.6	510	38	AA776918	AA776918 ac40e11.s
C 41	42.6	3.6	447	39	AA838639	AA838639 oe37h02.s
C 42	42.6	3.6	492	39	AA857507	AA857507 oh98e07.s
C 43	42.6	3.6	439	41	AI002243	AI002243 or72d12.s
44	42.2	3.5	377	23	D52591	D52591 HUM082C03B
45	42.2	3.5	342	23	D54105	D54105 HUM128H08B

# ALIGNMENTS

RESULT 1  
HSM008101  
ID HSM008101 standard; RNA; EST; 772 BP.  
XX  
AC AL043251;  
XX  
SV AL043251.1  
XX

```

NI e1401623
XX
XX 12-MAR-1999 (Rel. 59, Created)
DT 12-MAR-1999 (Rel. 59, Last updated, Version 1)
XX
XX Homo sapiens mRNA; EST DKFZp434K0223_r1 (from clone
DE DKFZp434K0223)
XX
XX EST: expressed sequence tag.
KW
XX
XX Homo sapiens (human)
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria;
OC Primates; Catarrhini; Homidae; Homo.
XX
XX [1]
RN 1-772
RP Blum H., Bauersachs S., Mewes W., Gassenhuber J., Wiemann S.;
RT Submitted (12-MAR-1999) to the EMBL/GenBank/DBJ databases.
RL MIPS, Am Klopferspitz 18a D-82152 Martinsried, GERMANY
XX
XX Clone from S. Wiemann, sequenced by LMU within the cDNA
CC sequencing consortium of the German Genome Project
CC No s1 sequence available
CC This clone is available at the RZPD in Berlin
CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
CC Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de
XX
XX Key Location/Qualifiers
FH
FH 1. 772
FT /db_xref="taxon:9606"
FT /organism="Homo sapiens"
FT /clone="DKFZp434K0223"
FT /clone_lib="434 (synonym: htes3). Vector pSport1; host
FT DH10B; sites NotI + SalI"
FT /dev_stage="adult"
FT /tissue_type="testis"
XX
XX Sequence 772 BP; 118 A; 250 C; 260 G; 143 T; 1 other;

Query Match 39.6%; Score 474.4; DB 54; Length 772;
Best Local Similarity 99.2%; Pred. No. 2.6e-98;
Matches 497; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

QY 1 GATGCTGACATGAGGACATTTTGTCTGCAAGTCCGCTATGCGCTGGGATGCAG 60
DB 256 GATGCTGACATGAGGACATTTTGTCTGCAAGTCCGCTATGCGCTGGGATGCAG 315
QY 61 GACCGGACCATCCAGACAGTGCATCTCTGCTCCAGCTCTGCTGAGATTCACACTGCC 120
DB 316 GACCGGACCATCCAGACAGTGCATCTCTGCTCCAGCTCTGCTGAGATTCACACTGCC 375
QY 121 GCCCGCCACAGCAGGTTGGAGACAGTGCAGGGGATGCGGCTGTCGCCCGCAGGGTCG 180
DB 376 GCCCGCCACAGCAGGTTGGAGACAGTGCAGGGGATGCGGCTGTCGCCCGCAGGGTCG 435
QY 181 GTGTTTCCCAAGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 240
DB 436 GTGTTTCCCAAGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 495
QY 241 CTGGTGGGACCCAGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 300
DB 496 CTGGTGGGACCCAGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 555
QY 301 CGGCTCGTTTACTCCCGGATGTTCCGCGCTGATGGGCTGGAAGACCGCTGGGGTCAG 360
DB 556 CGGCTCGTTTACTCCCGGATGTTCCGCGCTGATGGGCTGGAAGACCGCTGGGGTCAG 615
QY 361 GAGGTGATCTCAGCAATGAGGACCTGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 420
DB 616 GAGGTGATCTCAGCAATGAGGACCTGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 674

```

```

QY 421 ATGGTTGCCGACTGGTTCGCTTCTACCCCGGGTGCACCGGGTGCATGAGGCTGTGCTG 480
DB 675 ATGGTTGCCGACTGGTTCGCTTCTACCCCGGGTGCACCGGGTGCATGAGGCTGTGCTG 734
QY 481 CGGGTAGAGCTCTATGGGTGC 501
DB 735 CGGGTAGA-CTCTATGGGTGC 754

RESULT 2
AA073943
LOCUS
DEFINITION
mm97h12.r1 Stratagene mouse heart (#937316) Mus musculus cDNA clone
IMAGE:536423 5' similar to TR:G1161063 G1161063 RECEPTOR KINASE. ;
mRNA sequence.
AA073943
NID 91595690
VERSION AA073943.1 GI:1595690
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 433)
AUTHORS Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT On Sep 12, 1996 this sequence version replaced gi:1400901.

Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:323359
Possible reversed clone: similarity on wrong strand
Seq primer: -28ml3 rev1 ET from AmerSham
High quality sequence stop: 429.
FEATURES
source
1. .433
/organism="Mus musculus"
/strain="NIH/Swiss"
/db_xref="taxon:10090"
/clone="IMAGE:536423"
/clone_lib="Stratagene mouse heart (#937316)"
/sex="pooled"
/tissue_type="heart"
/dev_stage="13 day embryos"
/lab_host="SOLR (kanamycin resistant)"
/notes="Organ: heart; Vector: pBluescript SK-; Site:1:
EcoRI; Site:2: XhoI; Cloned unidirectionally. Primer:
Oligo dT, 93 pooled NIH/Swiss 13 day embryo hearts.
Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5'
adaptor sequence: 5' GAATTCGGCAGCAG 3' -3' adaptor
sequence: 5' CTCGAGTTTTTTTTTTT 3'"
BASE COUNT 78 a 103 c 155 g 96 t . 1 others
ORIGIN

Query Match 26.3%; Score 314.8; DB 28; Length 433;
Best Local Similarity 84.7%; Pred. No. 3.7e-62;
Matches 365; Conservative 0; Mismatches 63; Indels 3; Gaps 1;

QY 294 GAGCTACCGGCTGGTACTCCCGGGATGTCGCCGCTGGTGGCTGGAAGACCGCTG 353

```

```

Db 1 GAACTATCGTTCGGTACTCCGAGATGCCCGCCCTGGATGACTGGAAGACCGCTG 60
QY 354 GGGTCAGGAGGTGATCTCAGGCAATGAGGACCTGAGGAGGTGGTCTCAAGGACCTTGG 413
Db 61 GGGACAGGAGGTGATTTGGGTAAAGAGATCCCGGGGAGTAGTCTGAAGGACCTTGG 120
QY 414 GCGCCCATGTTGGCCGACTGGTTCGCTTCTACCCCGGGGTGACCGGGTCAATGAGCT 473
Db 121 ACCCGCATGTTGGCGGCTGGTCCGCTTCTACCCCGGGGTGACCGGATCATGAGT 180
QY 474 CTGCTGCGGGTAGACTCTATGGCTGCCTCTGAGGAGGATGACCTCTCTTACACGCG 533
Db 181 CTGCTTTCGGGGTGGAGCTCTATGGCTGCCTCTGGCGGATGGAGCTCTCTCATACAGC 240
QY 534 CCTGTGGGCGACACATATTTATCTGAGG---CCGTGTAACCTCAAGACTCCACCTA 590
Db 241 CCCCGTGGAGNAGCATGCACTATCTGAGGTGATGGTACATCTCAATGATCCACTTA 300
QY 591 TGACGACATACCGTGGCGGAGCTGAGTATGAGGTATGGGGTCTGGCCAGCTGGGT 650
Db 301 CGATGATATACTGCTGGAGGGCTGCACTATGCGGTCTGGCCAGCTGGCAGATGGCGT 360
QY 651 GGTGGGCTGGATGACTTTAGGAGACTCAGGAGCTGCGGGTCTGGCCAGCTGATGACTA 710
Db 361 GGTGGGCTGGATGATTTAGGAGAGCTCAGGAGCTGCGGGTCTGGCCAGCTGATGACTA 420
QY 711 TGTGGGATGGA 721
Db 421 TGTGGGATGGA 431

```

```

RESULT 3
AA622035/c
LOCUS
DEFINITION
nq23b08.s1 NCI CGAP Col0 Homo sapiens cDNA clone IMAGE:1144695 3'
similar to TR:G1160925 G1160925 RECEPTOR KINASE. ; mRNA sequence.
ACCESSION
AA622035
NID
92525911
VERSION
AA622035.1 GI:2525911
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 568)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
On Sep 12, 1996 this sequence version replaced gi:11405009.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck,
M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.11nl.gov/bbrp/image/image.html
Insert Length: 1405 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 423.
FEATURES
Location/Qualifiers
1..568
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="3"
/clone="IMAGE:1144695"

```

```

/clone_lib="NCI_CGAP_Col0"
/tissue_type="colon tumor RER+"
/lab_host="DH10B"
/notes="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
RER+ colon tumor, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7T3
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaudo (N-Soares4)."
BASE COUNT 150 a 158 c 131 g 127 t 2 others
ORIGIN

```

```

Query Match 17.68; Score 210.8; DB 36; Length 568;
Best Local Similarity 95.88; Pred. No. 1.7e-38;
Matches 226; Conservative 0; Mismatches 9; Indels 1; Gaps 1;
QY 502 CTCTGGAGGATGACTCTCTTACACCCGCCCTGTCGGGAGACAAATGTTATCT 561
Db 568 CTCTGGAGGATGACTCTCTTACNATCTGCTGT-GGGCAGACAAATGTTATCT 510
QY 562 GAGCCGCTGACCTCAACGACTCCACCTATGAGGACATACCTGGCGGACTGCAGTAT 621
Db 509 GAGCCGCTGACCTCAACGACTCCACCTATGAGTACATACCTGGCGGACTGCAGTAT 450
QY 622 GGGGCTGCGCCAGCTGGCAGATGGTGGTGGGCTGGATGACTTTAGGAAGTCTAG 681
Db 449 GGGGCTGCGCCAGCTGGCAGATGGTGGTGGGCTGGATGACTTTAGGAAGTCTAG 390
QY 682 GAGCTGCGGGTCTGGCCAGCTATGACTATGTGGATGGAGCAACACAGCTTCTC 737
Db 389 GAGCTGCGGGTCTGGCCAGCTATGACTATGTGGATGGAGCAACACAGCTTCTC 334

```

```

RESULT 4
ID HSM011191 standard; RNA; EST; 257 BP.
XX AC AL046341;
XX SV AL046341.1
XX NI e1404815
XX DT 12-MAR-1999 (Rel. 59, Created)
XX DT 12-MAR-1999 (Rel. 59, Last updated, Version 1)
XX DE Homo sapiens mRNA; EST DKFZp434J247_r1 (from clone
XX DE DKFZp434J247)
XX EST; expressed sequence tag.
XX KW
XX OS Homo sapiens (human)
XX OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria;
XX OC Primates; Catarrhini; Hominidae; Homo.
XX [1]
XX RP 1-257
XX RA Koehler K., Beyer A., Mewes W., Gassenhuber J., Wiemann S.;
XX RT Submitted (12-MAR-1999) to the EMBL/GenBank/DBJ databases.
XX RL MIPS, Am Klopferspitz 18a D-82152 Martinsried, GERMANY
XX CC Clone from S. Wiemann, sequenced by BMFZ within the cDNA
XX CC sequencing consortium of the German Genome Project
XX CC No s1 sequence available
XX CC This clone is available at the RZPD in Berlin
XX CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
XX CC Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de
XX FH
Key Location/Qualifiers

```

```

FH source
FT 1. .257
FT /db_xref="taxon:9606"
FT /organism="Homo sapiens"
FT /clone="DKFZp43J247"
FT /clone_lib="434 (synonym: htes3). Vector pSport1; host
FT DH10B; sites NotI + SalI"
FT /dev_stage="adult"
FT /tissue_type="testis"
FT
XX Sequence 257 BP; 40 A; 97 C; 67 G; 51 T; 2 other;

Query Match 14.9%; Score 177.8; DB 54; Length 257;
Best Local Similarity 97.8%; Pred. No. 3.9e-31;
Matches 179; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1015 TACTCTTCAGCAATCTCTTCATCTCTGATGTGTGACAAATCTCTTCGGGCACTG 1074
DB 1 TTAATCTTCAGCAATCTCTCTTCATCTCTGATGTGTGGAACAATTCNCGTCCGGCACTG 60

QY 1075 GGAGGCACCTTCCCGCCAGCCCTGTGTGGCCGCTGCGCCACCTCCACCACTTGAGC 1134
DB 61 GGAGGCACCTTCCCGCCAGCCCTGTGTGGCCGCTGCGCCACCTCCACCACTTGAGC 120

QY 1135 AGCTTGGAGCTGGAGCCAGAGCCAGAGCCCGTGGCCAGCCCGAGGGGAGCCCGACC 1194
DB 121 AGCTTGGAGCTGGAGCCAGAGCCAGAGCCCGTGGCCAGCCCGAGGGGAGCCCGACC 180

QY 1195 GCC 1197
DB 181 GCC 183

RESULT 5
LOCUS AI300297/c 245 bp mRNA EST 29-JAN-1999
DEFINITION qm59h12.x1 Soares_placenta_8to9weeks_2NbHP8to9w Homo sapiens CDNA
clone IMAGE:1894055 3' similar to SW:EDD1.HUMAN Q08345 EPITHELIAL
DISCOIDIN DOMAIN RECEPTOR 1 PRECURSOR ; mRNA sequence.
ACCESSION AI300297
NID 93959643
VERSION AI300297.1 GI:3959643
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 245)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Aug 21, 1998 this sequence version replaced.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 921 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 59.
Location/Qualifiers
1..245
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1894055"
/clone_lib="Soares_placenta_8to9weeks_2NbHP8to9w"
/dev_stage="two placentae; one from 8 weeks and another
from 9 weeks post conception"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Placenta; Vector: pT73D (Pharmacia) with a

modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGGCGCGCGATTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library constructed by Bento Soares and
M.Fatima Bonaldo."

BASE COUNT 58 a 48 c 61 g 78 t
ORIGIN

Query Match 10.0%; Score 119.4; DB 44; Length 245;
Best Local Similarity 95.3%; Pred. No. 7.1e-18;
Matches 123; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 ATGCTGACATGAAGGACATTTTGATCTGCAAGTGGCGCTATGCTGCGCATGCGAGG 61
DB 189 ATGCTGACATGAAGGACATTTTGATCTGCAAGTGGCGCTATGCTGCGCATGCGAGG 130

QY 62 ACCGGACCATCCACAGTGTGATCTCTGCTTCCAGCTCCTGTGTCAGATTCACATGCGG 121
DB 129 ACCGGACCATCCACAGTGTGATCTCTGCTTCCAGCTCCTGTGTCAGATTCACATGCGG 70

QY 122 CCCGCCACA 130
DB 69 CTCGCCCA 61

RESULT 6
LOCUS R85358 235 bp mRNA EST 14-AUG-1995
DEFINITION yo39h12.r1 Soares adult brain N2b4HB55Y Homo sapiens CDNA clone
IMAGE:180359 5' similar to SP:S37402 S37402 TRKE PROTEIN - ; mRNA
sequence.
ACCESSION R85358
NID 9943764
VERSION R85358.1 GI:943764
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 235)
AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevisan, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
Insert Size: 2589
High quality sequence starts: 1
High quality sequence stops: 1
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Insert Length: 2589 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 1.
Location/Qualifiers
1..235
/organism="Homo sapiens"
/db_xref="GDB:3827004"

FEATURES
Source
```









Db 159 AGTCACACTGGAAGTGCACAACTGGCTACTGACAAAGATGATCCCACTTATGTGACT 218  
 Qy 746 ATGTGGAGATGGAGTTGAGTTGACCGGCTGAGGGCTTCCAGGCTATGACGTCCACT 805  
 Db 219 TCATTGAGACGTATATCTAGTGAACAGGTGAATCCATTAAGAAGTGGGTGACACAG 278  
 Qy 806 GTAAACATGCACAGCTGGGAGCCGCTGCTGCTGCGGGGTGGAAATGTCGCTT 860  
 Db 279 TGACCAACTACGAAGTGGTCCCTGAGCTGCAATGGCAGATATCTCTT 333

## RESULT 12

AA597143 503 bp mRNA EST 19-SEP-1997  
 LOCUS V035f03.r1 Barstead mouse irradiated colon MPLRB7 Mus musculus CDNA  
 DEFINITION clone IMAGE:1051901 5' similar to gb:xi2812 Murine mRNA for  
 macrophage ferritin heavy subunit (MOUSE);, mRNA sequence.

## ACCESSION

AA597143

## NID

92412578

## VERSION

AA597143.1

## KEYWORDS

GI:2412578

## SOURCE

EST.

## ORGANISM

house mouse.

## REFERENCE

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 503)

Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,

Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,

Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,

Theising,B., Wylie,T., Lennon,G., Soares,B., Wilton,R. and

Waterston,R.

The WashU-HMI Mouse EST Project

Unpublished (1996)

On Sep 12, 1996 this sequence version replaced gi:1324621.

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:583477

Putative full length read

vector to vector length is 508

Seq primer: -28m13 rev2 ET from Amersham

High quality sequence stop: 94.

Location/Qualifiers

1. 503

/organism="Mus musculus"

/strain="FVB/N"

/db\_xref="taxon:10090"

/map="1"

/clone="IMAGE:1051901"

/dev\_stage="8 weeks"

/lab\_host="DH10B"

/note="Vector: p7T3D-Pac (Pharmacia) with a modified

polylinker; Site\_1: EcoRI; Site\_2: NotI; Tissue obtained

from 8 week old mouse. Colon was harvested 72 hours after

irradiation with 1400 Gys. 1st strand cDNA was primed

with a Not I - oligo(dT) primer

[5'-GTTCAGATCTGAGTGGAGCGCCGCTTTTCTTTTCTTTTCTTTT

T 3']; double-stranded cDNA was ligated to Eco RI

adaptors [AATTCGGATCCTTG], digested with Not I and cloned

into the Not I and Eco RI sites of the modified p7T3

vector. Library constructed by Bob Barstead.

144 a 115 c 122 g 122 t

BASE COUNT

ORIGIN

Query Match 3.7%; Score 44.6; DB 35; Length 503;  
 Best Local Similarity 49.4%; Pred. No. 0.89;  
 Matches 116; Conservative 0; Mismatches 119; Indels 0; Gaps 0;  
 QY 626 GTCTGGCCAGCTGCCAGATGGTGTGGTGGCGCTTGGAGTCTTGGAGAGAGTCCAGGAGC 685  
 Db 81 GACTGGGAGCGGGCTGAATGCAATGAGTGTGCTGCTGCTTGGAAAAGAGTGTGAATC 140  
 QY 686 TCGGGGTCTGGCCAGGCTATGACTATGTGGGATGGAGCAACACAGCTTCTCCAGTGGCT 745  
 Db 141 AGTCACTACTGGAAGTGCACAAAGTGGCTACTGACAAAGATGATCCCACTATGTGACT 200  
 QY 746 ATGTGGAGATGGAGTTTGGAGTTTGACCGGCTGAGGGCTTCCAGGCTATGAGAGTCCACT 805  
 Db 201 TCATTGAGAGCTATTATCTGAGTGAACAGGTGAATCCATTAAAGAACTGGGTGACACG 260  
 QY 806 GTAACAACATGCACAGCTGGGAGCCGCTCTGCTGCGGGGTGGAAATGTCGCTT 860  
 Db 261 TGACCAACTTACCAAGATGGTGGCTGCCCTGAAGCTGGCATGGCAGAAATATCTCTT 315

## RESULT 13

AA691209

LOCUS

DEFINITION

AA691209 484 bp mRNA EST 16-DEC-1997

vt34e06.r1 Barstead mouse proximal colon MPLRB6 Mus musculus CDNA

clone IMAGE:1164994 5' similar to gb:xi2812 Murine mRNA for

macrophage ferritin heavy subunit (MOUSE);, mRNA sequence.

AA691209

NID

92892145

VERSION

AA691209.1

KEYWORDS

EST.

SOURCE

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;

Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 484)

Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,

Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,

Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,

Theising,B., Wylie,T., Lennon,G., Soares,B., Wilton,R. and

Waterston,R.

The WashU-HMI Mouse EST Project

Unpublished (1996)

On Sep 12, 1996 this sequence version replaced gi:1397754.

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:630906

Seq primer: -28m13 rev2 ET from Amersham

High quality sequence stop: 357.

Location/Qualifiers

1. 484

/organism="Mus musculus"

/strain="FVB/N"

/db\_xref="taxon:10090"

/clone="IMAGE:1164994"

/dev\_stage="7 day juvenile"

/lab\_host="DH10B"

/note="Vector: p7T3D-Pac (Pharmacia) with a modified

polylinker; Site\_1: EcoRI; Site\_2: NotI; 1st strand cDNA

was primed with a Not I - oligo(dT) primer [5'

TGTTACGATCTGAAGTGGAGCGCCGCTTTTCTTTTCTTTTCTTTT

3']; double-stranded cDNA was ligated to Eco RI adaptors

[AATTCGGATCCTTG], digested with Not I and cloned into the

Not I and Eco RI sites of the modified pT73 vector.  
Library constructed by Bob Barstead. "

BASE COUNT  
ORIGIN

139 a 109 c 116 g 120 t

Query Match 3.7%; Score 44.6; DB 37; Length 484;  
Best Local Similarity 49.4%; Pred. No. 0.87;  
Matches 116; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

QY 626 GTCTGGCCAGCTGCACATGGTGTGGTGGCTGGATGACCTTTAGGAAGAGTCAGGAGC 685  
DB 65 GACTGGGAGAGCGGCTGAATGCAATGAGTGTGCACCTTGGAAAGAGTGTGAATC 124

QY 686 TCGGGTCTGGCCAGGCTATGATGTGGATGGAGCAACACAGCTTCTCCAGTGGCT 745  
DB 125 AGTCACTACTGGAATGCACAACTGGCTACTGACAAAGATGATCCCATTTGTGACT 184

QY 746 ATGTGGAGATGGAGTTTGAAGTTGACCGGCTGAGGGCCCTCCAGGCTATGAGGTCCACT 805  
DB 185 TCATTGAGACGTATTATCTGAGTGACAGGTTGAATCCATTAAAGAACTGGGTGACCAAG 244

QY 806 GTAAACAATGACACAGCTGGGAGCCCTGCTGCTGGCGGGGTGGAATGTCGCTT 860  
DB 245 TGACCAACTTACCAAGATGGGTGCCCTGGAAGCTGGCATGGCAGATATCTCTT 299

RESULT 14  
AA691386

LOCUS AA691386 507 bp mRNA EST 16-DEC-1997  
DEFINITION vU5a10.r1 Barstead mouse myotubes MPLRB5 Mus musculus cDNA clone  
IMAGE:1180698 5', similar to gb:X12812 Murine mRNA for macrophage  
ferritin heavy subunit (MOUSE);, mRNA sequence.

ACCESSION  
NID

AA691386.1 GI:2692322

VERSION  
KEYWORDS

SOURCE  
ORGANISM

house mouse.  
Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
AUTHORS

1 (bases 1 to 507)  
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,  
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and  
Waterston, R.

TITLE  
JOURNAL

COMMENT

The WashU-HHMI Mouse EST Project  
Unpublished (1996)  
On Sep 12, 1996 this sequence version replaced gi:1402308.

Contact: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project

Washington University School of MedicineP

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@wustl.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:638546

Seq primer: -28ml3 rev2 ET from Amersham

High quality sequence stop: 423.

Location/Qualifiers

1..507

/organism="Mus musculus"

/strain="C3H"

/db\_xref="taxon:10090"

/map="925E08; 2; 2q12.3-2q14.3"

/clone="IMAGE:1180698"

/clone\_lib="Barstead mouse myotubes MPLRB5"

/cell\_line="C2C12"

/lab\_host="DH10B"

FEATURES  
source

FEATURES

/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: EcoRI; Site\_2: NotI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5'  
TGTTACGAATCTGAAGTGGAGCGGCCCTTTTTTTTTTTTTTTTTTTT  
3']; double-stranded cDNA was ligated to Eco RI adaptors  
[AATTCGATCCTTG], digested with Not I and cloned into the  
Not I and Eco RI sites of the modified pT73 vector.  
Library constructed by Bob Barstead. The C2C12 cell line  
(available from ATCC, catalog # CRL-1772) differentiates  
rapidly, forming contractile myotubes and producing  
characteristic muscle proteins. "

BASE COUNT  
ORIGIN

144 a 113 c 128 g 122 t

Query Match 3.7%; Score 44.6; DB 37; Length 507;

Best Local Similarity 49.4%; Pred. No. 0.89;

Matches 116; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

QY 626 GTCTGGCCAGCTGCACATGGTGTGGTGGCTGGATGACCTTTAGGAAGAGTCAGGAGC 685

DB 87 GACTGGGAGAGCGGCTGAATGCAATGAGTGTGCACCTTGGAAAAGAGTGTGAATC 146

QY 686 TCGGGTCTGGCCAGGCTATGACTATGTGGATGGAGCAACACAGCTTCTCCAGTGGCT 745

DB 147 AGTCACTACTGGAACCTGCACAACTGGCTACTGACAAAGATGATCCCATCTATGTGACT 206

QY 746 ATGTGGAGATGGAGTTTGAAGTTGACCGGCTGAGGGCCCTCCAGGCTATCCAGGTCCACT 805

DB 207 TCATTGAGACGTATTATCTGAGTGAACAGGTGAATCCATTAAAGAACTGGGTGACCAAG 266

QY 806 GTAAACAATGACACAGCTGGGAGCCCTGCTGCTGGCGGGGTGGAATGTCGCTT 860

DB 267 TGACCAACTTACGCAAGATGGGTGCCCTGGAAGCTGGCATGGCAGATATCTCTT 321

RESULT 15  
AA692234

LOCUS  
DEFINITION

AA692234 518 bp mRNA EST 16-DEC-1997  
vU23H01.r1 Barstead mouse myotubes MPLRB5 Mus musculus cDNA clone  
IMAGE:1163953 5', similar to gb:J03941 Mouse ferritin heavy chain  
(MOUSE);, mRNA sequence.

ACCESSION  
NID

AA692234

VERSION

AA692234.1 GI:26931172

KEYWORDS

SOURCE

ORGANISM

house mouse.  
Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;  
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE  
AUTHORS

1 (bases 1 to 518)  
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,  
Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,  
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,  
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and  
Waterston, R.

TITLE

JOURNAL

COMMENT

The WashU-HHMI Mouse EST Project

Unpublished (1996)

On Sep 19, 1997 this sequence version replaced gi:1517401.

Contact: Marra M/Mouse EST Project

WashU-HHMI Mouse EST Project

Washington University School

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@wustl.wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:629865

Seq primer: -28ml3 rev2 ET from Amersham

High quality sequence stop: 403.

Location/Qualifiers





Db 196 GACCGGACCATCCACAGACAGTACATCTGCTTCCAGCTCCTGGTGCAGATTCACATGCC 255  
Qy 121 GCCGCCACACAGAGTTGGAGAGCAGTACAGGGGATGGGCTGGTGGCCCGCAGAGGTCG 180  
Db 256 GCCGCCACACAGAGTTGGAGAGCAGTACAGGGGATGGGCTGGTGGCCCGCAGAGGTCG 315  
Qy 181 GTGTTTCCACAGAGGAGGAGTACTTGCAGGTGATCTACACAGCTGCACCTGGTGGCT 240  
Db 316 GTGTTTCCACAGAGGAGGAGTACTTGCAGGTGATCTACACAGCTGCACCTGGTGGCT 375  
Qy 241 CTGTGGGACCCACAGGACGATCCGCGGGCTGGCAAGAGTCTCCCGAGAGTAC 300  
Db 376 CTGTGGGACCCACAGGACGATCCGCGGGCTGGCAAGAGTCTCCCGAGAGTAC 435  
Qy 301 CGGTGCTTACTCCCGGATGGTGGCGCTGGATGGCTGGAGGACCGCTGGGGTCAG 360  
Db 436 CGGTGCTTACTCCCGGATGGTGGCGCTGGATGGCTGGAGGACCGCTGGGGTCAG 495  
Qy 361 GAGGTGATCTCAGGAATGAGGACCTGAGGAGTGGTGTGAAGGACCTTGGGCCCCC 420  
Db 496 GAGGTGATCTCAGGAATGAGGACCTGAGGAGTGGTGTGAAGGACCTTGGGCCCCC 555  
Qy 421 ATGTTGCCCGACTGGTTCGCTTACCCCGGCTGACCGGTCATGAGCTGTGCTG 480  
Db 556 ATGTTGCCCGACTGGTTCGCTTACCCCGGCTGACCGGTCATGAGCTGTGCTG 615  
Qy 481 CGGTGAGCTCTATGGTGGCTCTGGAGGATGGATCTCTTACACGGCCCTGTG 540  
Db 616 CGGTGAGCTCTATGGTGGCTCTGGAGGATGGATCTCTTACACGGCCCTGTG 675  
Qy 541 GGGCAGACAATGTTATCTAGAGCGCTGACCTCAACGACTCCACCTATGACGGACAT 600  
Db 676 GGGCAGACAATGTTATCTAGAGCGCTGACCTCAACGACTCCACCTATGACGGACAT 735  
Qy 601 ACCGTGGCGGACTGCAGTATGGGGCTGGCCAGCTGGCAGATGTGTGGGGCTG 660  
Db 736 ACCGTGGCGGACTGCAGTATGGGGCTGGCCAGCTGGCAGATGTGTGGGGCTG 795  
Qy 661 GATGACTTTAGGAAGTACAGAGCTGGGCTGGCCAGCTATGACTATGTGGGATG 720  
Db 796 GATGACTTTAGGAAGTACAGAGCTGGGCTGGCCAGCTATGACTATGTGGGATG 855  
Qy 721 AGCAACACAGCTTCTCAGTGGCTATGTGAGATGAGATTTGATTTGACCGGCTAGG 780  
Db 856 AGCAACACAGCTTCTCAGTGGCTATGTGAGATGAGATTTGATTTGACCGGCTAGG 915  
Qy 781 GCCTTCCAGGCTATGCAGGTCCACTGTAAACAATGCACAGCTGGGAGCCGCTGCCT 840  
Db 916 GCCTTCCAGGCTATGCAGGTCCACTGTAAACAATGCACAGCTGGGAGCCGCTGCCT 975  
Qy 841 GCGGGGTGGAATGTCGCTTCCGGCGTGGCCCTGCCATGGCTGGGAGGGGAGCCCATG 900  
Db 976 GCGGGGTGGAATGTCGCTTCCGGCGTGGCCCTGCCATGGCTGGGAGGGAGCCCATG 1035  
Qy 901 GCCCACACCTTAGGGGGAACCTGGGGAGCCCAAGAGCCCGGCTGTCTAGTCCGCCCT 960  
Db 1036 GCCCACACCTTAGGGGGAACCTGGGGAGCCCAAGAGCCCGGCTGTCTAGTCCGCCCT 1095  
Qy 961 GCGGCCGTGTGGCTGCTTCTGAGTGGCGCTTCCCTTTTGGGGCCCTGGTTACTC 1020  
Db 1096 GCGGCCGTGTGGCTGCTTCTGAGTGGCGCTTCCCTTTTGGGGCCCTGGTTACTC 1155  
Qy 1021 TTCAGCGAAATCTCCTTCTATCTGATGTGTGAACAATTCCTTCGGGCACTGGGAGG 1080  
Db 1156 TTCAGCGAAATCTCCTTCTATCTGATGTGTGAACAATTCCTTCGGGCACTGGGAGG 1215  
Qy 1081 ACCTTCCGCGACGCCCTTGGTGGCCGCTGGCCACCTCCACACATTCAGCAGCTTG 1140  
Db 1216 ACCTTCCGCGACGCCCTTGGTGGCCGCTGGCCACCTCCACACATTCAGCAGCTTG 1275  
Qy 1141 GAGCTGGAGCCACAGGACGACGAGCCCTGGCCAGCCGAGGGGAGCCGACCGCC 1197

Db 1276 GAGCTGGAGCCAGAGCCAGCAGCCCTGGCCAGCCCGAGGGAGCCGACCGCC 1332

## RESULT 2

US-08-445-640-7  
; Sequence 7, Application US/08445640  
; Patent No. 5709858  
; GENERAL INFORMATION:  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Mark, Melanie R.  
; APPLICANT: Scadden, David T.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Baron, Will F.  
; TITLE OF INVENTION: Protein Tyrosine Kinases  
; NUMBER OF SEQUENCES: 35  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/445,640  
; FILING DATE: 22-MAY-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/170558  
; FILING DATE: 20-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/157563  
; FILING DATE: 23-NOV-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hasak, Janet E.  
; REGISTRATION NUMBER: 28,616  
; REFERENCE/DOCKET NUMBER: 854C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-1896  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1197 bases  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-445-640-7

Query Match 100.0%; Score 1197; DB 2; Length 1197;  
Best Local Similarity 100.0%; Pred. No. 2e-287;  
Matches 1197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATGCTGACATCAAGGACATTTTGATCTGCAAGTCCGCTATGCCCTGGGCATGCAG 60  
Db 1 GATGCTGACATCAAGGACATTTTGATCTGCAAGTCCGCTATGCCCTGGGCATGCAG 60  
Qy 61 GACCGGACCATCCACAGACAGTGCATCTGTCTCCAGCTCCTGTGTCAAGTTCACCTGCC 120  
Db 61 GACCGGACCATCCACAGACAGTGCATCTGTCTCCAGCTCCTGTGTCAAGTTCACCTGCC 120  
Qy 121 GCCGCCACAGCAGGTGGAGAGCAGTGCAGGGGATGGGCTGGTGGCCCGCAGGTCG 180  
Db 121 GCCGCCACAGCAGGTGGAGAGCAGTGCAGGGGATGGGCTGGTGGCCCGCAGGTCG 180  
Qy 181 GTGTTTCCCAAGGAGGAGGAGTACTTGCAGGTGGATCTACACAGCTGCACCTGGTGGCT 240  
Db 181 GTGTTTCCCAAGGAGGAGGAGTACTTGCAGGTGGATCTACACAGCTGCACCTGGTGGCT 240



QY 241 CTGGTGGGACACCGAGCGGATGCGGGGGGCTGGGCAAGAGTTCTCCGGAGCTAC 300  
DB 241 CTGGTGGGACACCGAGCGGATGCGGGGGGCTGGGCAAGAGTTCTCCGGAGCTAC 300  
QY 301 CGGCTGGTACTCCCGGAGTGTCCCGCTGGATGGGCTGGAAGACCGCTGGGGTAC 360  
DB 301 CGGCTGGTACTCCCGGAGTGTCCCGCTGGATGGGCTGGAAGACCGCTGGGGTAC 360  
QY 361 GAGGTGATCTCAGCAATGAGGACCTGAGGAGTGTCTGAAAGACCTTGGGCCCCC 420  
DB 361 GAGGTGATCTCAGCAATGAGGACCTGAGGAGTGTCTGAAAGACCTTGGGCCCCC 420  
QY 421 ATGGTTCCCGGAGTGTCTGCTTCTACCCCGGCTGACCGGGTCAAGAGCTGTGTG 480  
DB 421 ATGGTTCCCGGAGTGTCTGCTTCTACCCCGGCTGACCGGGTCAAGAGCTGTGTG 480  
QY 481 CGGCTAGAGCTCTATGGCTGCTCTGAGGAGTGTCTGCTTCTACACCGCCCTGTG 540  
DB 481 CGGCTAGAGCTCTATGGCTGCTCTGAGGAGTGTCTGCTTCTACACCGCCCTGTG 540  
QY 541 GGCAGACAATGATTTATCTGAGGCGGTGTACCTCAAGACTTCCACCTATGACGACAT 600  
DB 541 GGCAGACAATGATTTATCTGAGGCGGTGTACCTCAAGACTTCCACCTATGACGACAT 600  
QY 601 ACCGTGGGCGGACTGAGTATGGGGTCTGGGCGAGCTGGCAGATGGTGTGGGGTG 660  
DB 601 ACCGTGGGCGGACTGAGTATGGGGTCTGGGCGAGCTGGCAGATGGTGTGGGGTG 660  
QY 661 GATGACTTTAGGAAGAGTCAGAGCTGCGGCTGCGGAGGCTATGACTATGGGATGG 720  
DB 661 GATGACTTTAGGAAGAGTCAGAGCTGCGGCTGCGGAGGCTATGACTATGGGATGG 720  
QY 721 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGTGGAGTTTGAGTTTGACCGGCTGAG 780  
DB 721 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGTGGAGTTTGAGTTTGACCGGCTGAG 780  
QY 781 GCCTTCCAGGCTATGAGCTTCACTGTAAACATGACACGCTGGGAGCCGCTGTGCT 840  
DB 781 GCCTTCCAGGCTATGAGCTTCACTGTAAACATGACACGCTGGGAGCCGCTGTGCT 840  
QY 841 GCGGGGTGGAATGCTGCTTCCGGCTGCGCTGCGCATGGCTGGGAGGGGAGCCCATG 900  
DB 841 GCGGGGTGGAATGCTGCTTCCGGCTGCGCTGCGCATGGCTGGGAGGGGAGCCCATG 900  
QY 901 CGCCACAACCTAGGGGCAACCTGGGGACCCAGAGCCCGGCTGTCTCAGTGGCCCTT 960  
DB 901 CGCCACAACCTAGGGGCAACCTGGGGACCCAGAGCCCGGCTGTCTCAGTGGCCCTT 960  
QY 961 GCGGGCGTGTGGCTGCTTCTGAGTCCGCTTCTGAGTCCGCTTCTGCGGGGCTGTACTC 1020  
DB 961 GCGGGCGTGTGGCTGCTTCTGAGTCCGCTTCTGAGTCCGCTTCTGCGGGGCTGTACTC 1020  
QY 1021 TTCAGGAAATCTCTTCACTCTGATGTTGGAACAATTCCTTCCGGCACTGGGAGGC 1080  
DB 1021 TTCAGGAAATCTCTTCACTCTGATGTTGGAACAATTCCTTCCGGCACTGGGAGGC 1080  
QY 1081 ACTTCCCGCCAGCCGCTGTGGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1140  
DB 1081 ACTTCCCGCCAGCCGCTGTGGCGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1140  
QY 1141 GAGCTGGAGCCAGGCGGAGCCGCTGGCGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1197  
DB 1141 GAGCTGGAGCCAGGCGGAGCCGCTGGCGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1197

## RESULT 3

US-08-336-343A-1

: Sequence 1, Application US/08336343A

: Patent No. 567144

: GENERAL INFORMATION:

: APPLICANT: Ullrich, Axel

: APPLICANT: Alves, Frauke

; TITLE OF INVENTION: CCK-2, A No. 567714461 Receptor Tyrosine Kinase  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/336,343A  
; FILING DATE: 08-NOV-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7683-065  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-9741/8864  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3962 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: unknown  
; MOLECULE TYPE: cDNA  
; HYPOTHEetical: NO  
; ANTI-SENSE: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 321..3077  
; US-08-336-343A-1

Query Match 99.8%; Score 1192.2; DB 2; Length 3962;  
Best Local Similarity 99.7%; Pred. No. 4.3e-286;  
Matches 1194; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GATGCTGACATGAGGACATTTTGTATCTGCTGCAAGTGGCGCTATGCTTGGGCGATGCGAG 60  
DB 375 GATGCTGACATGAGGACATTTTGTATCTGCTGCAAGTGGCGCTATGCTTGGGCGATGCGAG 434  
QY 61 GACGGACCATCCAGACAGTGCATCTCTGCTTCCAGCTCTCTGCTGCTGCTGCTGCTGCTGCTG 120  
DB 435 GACGGACCATCCAGACAGTGCATCTCTGCTTCCAGCTCTCTGCTGCTGCTGCTGCTGCTGCTG 494  
QY 121 GCGCGCCACAGAGTTGGAGACAGTGCAGGGGATGGGGCTGGTGGCCCGCAGGCTCG 180  
DB 495 GCGCGCCACAGAGTTGGAGACAGTGCAGGGGATGGGGCTGGTGGCCCGCAGGCTCG 554  
QY 181 GTGTTTCCCAAGGAGGAGTACTTCCAGTGGATCTACAAGACTGCACCTTGGTGGCT 240  
DB 555 GTGTTTCCCAAGGAGGAGTACTTCCAGTGGATCTACAAGACTGCACCTTGGTGGCT 614  
QY 241 CTGTTGGGCAACCCAGGACCGGATGCGGGGCTGGGCAAGAGTTCTCCCGGAGCTAC 300  
DB 615 CTGTTGGGCAACCCAGGACCGGATGCGGGGCTGGGCAAGAGTTCTCCCGGAGCTAC 674  
QY 301 CGGCTGGCTTACTCCCGGATGCTGCCGCTGGATGGGTGAAGACCGCTGGGGTCTAG 360  
DB 675 CGGCTGGCTTACTCCCGGATGCTGCCGCTGGATGGGTGAAGACCGCTGGGGTCTAG 734  
QY 361 GAGTGTATCTCAGCAATGAGGACCTGAGGAGTGGTGTGAGGACCTTGGGCCCCCCC 420  
DB 735 GAGTGTATCTCAGCAATGAGGACCTGAGGAGTGGTGTGAGGACCTTGGGCCCCCCC 794

QY 421 ATGGTGGCCGACCTGGTTCGCTTCTACCCCGGGCTGACCGGGTCTGACGGTCTGCTG 480  
Db 795 ATGGTGGCCGACCTGGTTCGCTTCTACCCCGGGCTGACCGGGTCTGACGGTCTGCTG 854  
QY 481 CGGTAGAGCTCTATGGTGGCTCTGGAGGATGAGTCTGCTTACACCGCCCTGCTG 540  
Db 855 CGGTAGAGCTCTATGGTGGCTCTGGAGGATGAGTCTGCTTACACCGCCCTGCTG 914  
QY 541 GGGGAGCAATATGTTATCTAGGCGGCTGTACCTCAACGACTCCACTATGACGGACAT 600  
Db 915 GGGGAGCAATATGTTATCTAGGCGGCTGTACCTCAACGACTCCACTATGACGGACAT 974  
QY 601 ACCGTGGCGGACCTGACGATGGGGTCTGGCCAGCTGGCAGATGCTGTGGGGCTG 660  
Db 975 ACCGTGGCGGACCTGACGATGGGGTCTGGCCAGCTGGCAGATGCTGTGGGGCTG 1034  
QY 661 GATGACTTTAGGAAGTACGAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 720  
Db 1035 GATGACTTTAGGAAGTACGAGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1094  
QY 721 AGCAACACAGCTCTCCAGTGGCTATGTGGAGATGAGTTTGTAGTTTGTAGTTTGTAG 780  
Db 1095 AGCAACACAGCTCTCCAGTGGCTATGTGGAGATGAGTTTGTAGTTTGTAGTTTGTAG 1154  
QY 781 GCCTTCCAGGCTATGACGCTCCACTGTAAACATGACACGCTGGAGCCGCTGCTGCT 840  
Db 1155 GCCTTCCAGGCTATGACGCTCCACTGTAAACATGACACGCTGGAGCCGCTGCTGCT 1214  
QY 841 GCGGGGTGGAAATGTCGCTTCCGCGTGGCCCTGCCATGGCTGGGAGGGAGCCCATG 900  
Db 1215 GCGGGGTGGAAATGTCGCTTCCGCGTGGCCCTGCCATGGCTGGGAGGGAGCCCATG 1274  
QY 901 CGCCACAACCTAGGGGGCAACCTGGGGGACCCAGAGCCGGGCTGTCTCAGTGGCCCTT 960  
Db 1275 CGCCACAACCTAGGGGGCAACCTGGGGGACCCAGAGCCGGGCTGTCTCAGTGGCCCTT 1334  
QY 961 GCGGCGCTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
Db 1335 GCGGCGCTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1394  
QY 1021 TTCACGAAATCTCTTCTATCTGATGTGTGAACAAATCTCTCGGCACTGGGAGGC 1080  
Db 1395 TTCACGAAATCTCTTCTATCTGATGTGTGAACAAATCTCTCGGCACTGGGAGGC 1454  
QY 1081 ACCTTCCGCGGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140  
Db 1455 ACCTTCCGCGGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1514  
QY 1141 GAGCTGGAGCCGAGGCGGACGAGCCGCTGGCCAGCCCGGAGGAGCCGCGCC 1197  
Db 1515 GAGCTGGAGCCGAGGCGGACGAGCCGCTGGCCAGCCCGGAGGAGCCGCGCC 1571

## RESULT 4

US-08-336-343A-3  
; Sequence 3, Application US/08336343A  
; Patent No. 5677144  
; GENERAL INFORMATION:  
; APPLICANT: Ulrich, Axel  
; APPLICANT: Alves, Frauke  
; TITLE OF INVENTION: CCK-2, A No. 5677144el Receptor Tyrosine Kinase  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/336.343A  
; FILING DATE: 08-NOV-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7683-065  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-9741/8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3157 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: both  
; TOPOLOGY: unknown  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 370..2934  
; US-08-336-343A-3

Query Match 27.4%; Score 327.4; DB 2; Length 3157;

Best Local Similarity 58.8%; Pred. No. 2.2e-72;

Matches 644; Conservative 0; Mismatches 436; Indels 15; Gaps 4;

QY 13 AAGGACATTTTATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 72  
Db 433 AAGCTCAGGTTAATCCAGCTATATGCGCTATCTCTCTGCGCATGCTCAGAGGCGCAGATT 492  
QY 73 CCAGACAGTGCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 132  
Db 493 CCAGATGAGGACATCACAGCTTCCAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 552  
QY 133 AGGTTGGAGACAGTGAAGGAGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGG 192  
Db 553 AGGCTGGACTCAGAGAGAGGAGTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 612  
QY 193 GA---GGAGGAGTACTTGCAGTGGATCTTACAAACAGCTGACAGCTGCTGCTGCTGCTGCTG 249  
Db 613 GACCTGAAGGAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 672  
QY 250 ACCGAGGAGCGATGCGGGGGCTGGGCGAGGAGTCTCCCGGAGGCTACCGGCTGCTGCT 309  
Db 673 ACCGAGGAGCGATGCGAGGAGTCTGCGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 732  
QY 310 TACTCCGCGGATGGTGGCGCTGGATGGGCTGGAAGGACCGCTGGGGTCCAGGAGTGCATC 369  
Db 733 TACAGTCGGGATGGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 792  
QY 370 TCAGGCAATGAGGACCTGAGGAGTGGTGTGTAAGGACCTTGGGCGCCCGCCCGCTGCTGCTG 429  
Db 793 GATGAAATAGTAAACCCCTATGACATTTTCTTAAAGGACTTGGAGCCGCGCCATTTAGCC 852  
QY 430 CGACTGTTGCTGCTTACCCCGGCTGACCGGCTGACCGGCTGACGAGCTGCTGCTGCTGCTGCTG 489  
Db 853 AGATTTGCTCGGTTTATTCACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 912  
QY 490 CTCTATGCTGCTGCTGAGGAGTGGACTCTCTTACACCGCCCGCTGCTGCTGCTGCTGCTGCTG 549  
Db 913 CTTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 972  
QY 550 ATGTATTTATCTGAGGC-----CGTGTACCTCAAGGCTCCACCTATGACGAGCATACC 603  
Db 973 TTTGTACTCCCTGGAGGTTCCCATCATTTATCTGAATGATTTCTGTCTATGATGAGCTGTT 1032  
QY 604 GTGGCGGACTGCAGTATGCGGGTCTGGGCCAGCTGGCCAGATGGTGTGCTGCTGCTGCTGCTG 663

Db	1033	GGATACAGCATGACAGAAGGG--CTAGGSCCAATTGACCGATGCTGTCTGGCCTGGAC	1089
Qy	564	GACTTTTAGGAAGTGCAGGAGCTGGGGTCTGGCCAGCTATGACTATGTGGGATGGAGC	723
Db	1090	GATTTTCAACCCAGCCCATGAATACCACTGTGGCCGCTATGACTATGTGGCTGGCGG	1149
Qy	724	AACACAGCTTCTCCAGTGGCTATGTGGAGTGGAGTTTGAGTTTGACCGCTGAGGGCC	783
Db	1150	AACGAGAGTGCACCAATGCGTACATTGAGATCATGTTTGAATTGACCGCATCAGGAAT	1209
Qy	784	TTTCCAGGCTATGCAGGTCACCTGTAAACAATGCACAGCTGGGAGCCCGTCTGCCTGGC	843
Db	1210	TTCACTATCAATGAAGTCTCACTGCAACAACATGTTTGTCAAGGTGTGAAGATCTTTAAG	1269
Qy	844	GGGGTGGAAATGTCGCTTCGGCGTGGCCCTGCCATGGCTGGGAGGGGAGCCCATGCGC	903
Db	1270	GAGGTACAGTGTACTTTCGGCTCTGAA--GCCAGTGAGTGGGAACCTAATGCCATTTC	1326
Qy	904	CACAACCTAGGGGGCAACCTGGGGGACCCAGAGCCCGGGCTGTCTCAGTGTGCCCTTGGC	963
Db	1327	TTCCCCCTTGTCTCGATGACGTCAACCCCAAGTCTCGGGTTGTACGGGTGCCCTCCAC	1386
Qy	964	GGCGGTGTGGCTCGCTTTCTGCAAGTGC CGCCTCTCTTTTGGGGGGCCCTGGTTACTTTC	1023
Db	1387	CACCGAATGGCCAGTGCCATCAAGTGTCAATACCAATTTTGCAGATACCTCGATGATGTT	1446
Qy	1024	ACGGAATCTCCTTCACTCTGATGTGTGAACAATTCCTCTCCGGCACTGGGAGGCAC	1083
Db	1447	AGTGAGATCACCTTCCAATCAGATGCTGCAATGTACAACAACCTGAAGCCCTGCCAC	1506
Qy	1084	TTCCCCCAGCCCC	1098
Db	1507	TCCTCTATGGCACCC	1521

## RESULT

```

1  US-08-336-343A-5/C
2  Sequence 5, Application US/08336343A
3  Patent No. 5677144
4  GENERAL INFORMATION:
5  APPLICANT: Ullrich, Axel
6  APPLICANT: Alves, Frauke
7  TITLE OF INVENTION: CCK-2, A No. 5677144el Receptor Tyrosine Kinase
8  NUMBER OF SEQUENCES: 43
9  CORRESPONDENCE ADDRESS:
10 ADDRESSSEE: Pennie & Edmonds
11 STREET: 1155 Avenue of the Americas
12 CITY: New York
13 STATE: New York
14 COUNTRY: U.S.A.
15 ZIP: 10036-2711
16 COMPUTER READABLE FORM:
17 MEDIUM TYPE: Floppy disk
18 COMPUTER: IBM PC compatible
19 OPERATING SYSTEM: PC-DOS/MS-DOS
20 SOFTWARE: PatentIn Release #1.0, Version #1.30
21 CURRENT APPLICATION DATA:
22 APPLICATION NUMBER: US/08/336,343A
23 FILING DATE: 08-NOV-1994
24 CLASSIFICATION: 435
25 ATTORNEY/AGENT INFORMATION:
26 NAME: Coruzzi, Laura A.
27 REGISTRATION NUMBER: 30,742
28 REFERENCE/DOCKET NUMBER: 7683-065
29 TELECOMMUNICATION INFORMATION:
30 TELEPHONE: (212) 790-9090
31 TELEFAX: (212) 869-9741/8864
32 TELEX: 66141 PENNIE
33 INFORMATION FOR SEQ ID NO: 5:
34 SEQUENCE CHARACTERISTICS:
35 LENGTH: 3157 base pairs
36 TYPE: nucleic acid

```

Db 1831 TTCCCCCTTGCTGCGATGACGTCAACCCAGTCTCGGTTTGTACGGTGCCCTCTCCAC 1772  
Qy 964 GCGCGTGTGGCTCGCTTTCTTCAGTGGCGGCTTCCTCTTTTGGGGGCGCTGGTACTTTC 1023  
Db 1771 CACCAATGGCCAGTGCATCAAGTGTCAATACCATTTTGCAGATACCTGGATGATGTC 1712  
Qy 1024 AGCGAAATCTCCTTCATCTCTGATGTGGTGAACAATTCCTCTCCGGCACTGGGAGGCACC 1083  
Db 1711 AGTGAGATCACTTCCCAATCAGATGCTGCAATGTACAACAACCTCTGAAGCCCTGCCACC 1652  
Qy 1084 TTCCCGCCAGCCCC 1098  
Db 1651 TCTCCTATGGCACCC 1637

## RESULT 6

US-08-456-647B-19  
; Sequence 19, Application US/08456647B  
; Patent No. 5811516  
; GENERAL INFORMATION:  
; APPLICANT: Lemke Ph.D. et al., Greg E.  
; TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 4225 Executive Square, Suite 1400  
; CITY: La Jolla  
; STATE: CA  
; COUNTRY: US  
; ZIP: 92037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/456.647B  
; FILING DATE: 02-JUN-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/237,401  
; FILING DATE: 02-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/884,486  
; FILING DATE: 15-MAY-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Wetherell Ph.D., John R.  
; REGISTRATION NUMBER: 31,678  
; REFERENCE/DOCKET NUMBER: 07251/007002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 678-5070  
; TELEFAX: (619) 678-5099  
; INFORMATION FOR SEQ ID NO: 19:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3120 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
; IMMEDIATE SOURCE:  
; CLONE: Tyro-10  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 485..3047  
US-08-456-647B-19

Query Match 26.8%; Score 321; DB 3; Length 3120;  
Best Local Similarity 58.4%; Pred. No. 8.2e-71;  
Matches 640; Conservative 0; Mismatches 440; Indels 15; Gaps 4;

Qy 13 AAGGACATTTTGATCTCGCAAGTGGCGCTATGCCCTGGGCATGCAGGACCGGCACATC 72

Db 548 AAAGTCAGGTTAAATCCAGCCATATGCGGCTATCCTCTGGGCATGTCTAGAGGCGCACATT 607  
Qy 73 CCAGCAGATGACATCTCTGCTTCCAGCTCCTGGTCCAGATTCACATGCGCGCGCCACAGC 132  
Db 608 CCAGATGAGGACATCACAGCCTCAAGTCAGTGGTCAGAAATCCAGGCTGCCAATATGGG 567  
Qy 133 AGGTTGGAGAGCAGTGACGGGATGGGCGCTGGTCCCGGAGGCTCGGTTCGTTTCCCAAG 192  
Db 668 AGGCTGGACTCTGAAGAAGGAGATGGAGCTGGTCTCTGAGATTCAGATTCGAACCCGAT 727  
Qy 193 GA---GGAGGAGTACTTGCAGGTGATCTACAACGACTGCACCTGGTGGCTCTGGGCG 249  
Db 728 GACCTGAAGGAATTTCTGCAGATTGACTTGGGAACCCCTACACTTTATCACTCTTGTGGG 787  
Qy 250 ACCCAGGACGATGCGCGGGGCTGGGCAAGGAGTTCTCCCGGAGCTACCGGTGGCT 309  
Db 788 ACCCAGGCGGCATGCAGGGGCTATGGCATTGAATTTGACCCATGTACAAGATCAAC 847  
Qy 310 TACTCCCGGATGGTCCCGCTGGATGGGCTGGAAGGACCGCTGGGGTTCAGGAGGTGATC 369  
Db 848 TACAGTCGGGATGGCAGTCGCTGGATCTCTGGCGTAACCGCATGGGAAGCAGGTGCTT 907  
Qy 370 TCAGCAATGAGGACCCCTGAGGAGTGGTGTGAAGGACCTTGGCCCCCCTATGTTGCC 429  
Db 908 GATGGAACAGTAACCCCTTATGATGTATCTCTGAAGGACTTGGAGCCACCCATCGTCGCC 967  
Qy 430 CGACTGGTTCGCTTCTACCCCGGCTGACGGGCTCATGAGGCTCTGTCTCGGGTAGAG 489  
Db 968 AGATTTGTTCGCTTATCCAGTCACTGACCACCTCATGAAGCTGTGCATGAGGTTGAG 1027  
Qy 490 CTCTATGGCTCGCTCGAGGATGAGTCTCTGTCTTACACCGCCCTGTGGGCGACAGA 549  
Db 1028 CTTTATGGTTGTCTGGCTAGATGGCTTGGTATCTCTACAATGTCTCAGCTGAGCAGAG 1087  
Qy 550 ATGTAATTTATCTGAGC-----CGTGTACCTCAAGACTCCACCTATGAGGACATACC 603  
Db 1088 TTTGTACTCCCTGGAGGCTCCATCATTTATCTGAATGATTTCTGTCTATGATGGAGCTTT 1147  
Qy 604 GTGGCGGACTGCAGTATGGGGGCTGCGGCCAGCTGGCAGATGGTGTGGTGGGCTGGAT 663  
Db 1148 GGTACAGCATGACTGAAGGG---CTAGCCAGTTGACTGATGAGATATCGGGCTGAT 1204  
Qy 664 GACTTTAGGAAGATCAGGAGCTCGGGCTCTGGCCAGGCTATGACTATGTGGATGAGC 723  
Db 1205 GATTTTACCAGACCCATGAATACCACGCTGGCTGGCTGCTGACTACGTGGGATGGCG 1264  
Qy 724 AACCACAGCTTCTCCAGTGGCTATGTGAGATGGAGTTTGAGTTGACCGGCTGAGGSCC 783  
Db 1265 AACGAAAGTGCTACCAACGGTTTCATTGAGATCATGTTTGAATTTGACCGAATCAGSAA 1324  
Qy 784 TTCCAGGCTATGCAGTCCACTGTAAACAACATGCACAGCTGGGAGCGCGTCTGCTGGC 843  
Db 1325 TTTACTACCATGAAGTCCACTGCAACAACATGTTTGTGAAGGTGTGAAGATTTTAAAG 1384  
Qy 844 GGGGTGGAATGTGCTTCGCGCTGGCCCTGCCATGGCTGGAGGGGAGGCCATCGCC 903  
Db 1385 GAGGTCCAGTGTACTTTTCGCTCGGA---AGCCAGCAGTGGGAACCCACCTGCTGTAC 1441  
Qy 904 CACAACCTAGGGGGCAACCTGGGGGACCCAGAGCCCGGGCTGTCTCAGTCCCCCTTGGC 963  
Db 1442 TTTCCCTTGGTCTGTGACGATGTGAACCCCGAGTGGCCCGTTTGTACCGGTGCCCTCCAC 1501  
Qy 964 GCGCGTGTGGCTTCTCTGCAGTGGCGCTTCCCTCTTTTGGGGGCGCTGGTTACTCTTC 1023  
Db 1502 CACCGAATGGCCAGTGCATCAAGTGCACCAATACCATTTTGGCAGACAGTGGATGATTC 1561  
Qy 1024 AGCGAAATCTCCTTCATCTCTGATGTGGTGAACAATTCCTCTCCGGCACTGGAGGCACC 1083  
Db 1562 AGCAGATCACTTTTCCAAATCAGATGCTGCAATGTATAACAACCTCTGGAGCCCTTCCCACC 1621  
Qy 1084 TTCCCGCCAGCCCC 1098



COUNTRY: USA  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/670,707A  
FILING DATE: 26-JUN-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US94/13200  
FILING DATE: 15-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/212,133  
FILING DATE: 11-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/864,004  
FILING DATE: 07-APR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Greenlee, Lorraine L.  
REGISTRATION NUMBER: 27,894  
REFERENCE/DOCKET NUMBER: 75-95F  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 303/499-8080  
TELEFAX: 303/499-8089  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6402 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: not relevant  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
ORGANISM: Pig  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..6402  
US-08-670-707A-36

Query Match 3.9%; Score 46.6; DB 3; Length 6402;  
Best Local Similarity 45.9%; Pred. No. 0.0066;  
Matches 197; Conservative 0; Mismatches 229; Indels 3; Gaps 1;  
QY 22 TTGTATCTCTCCCAAGTGGCGCTATGCCCTGGCATGCGAGCCGACCATCCAGACAGT 81  
DB 5908 TGTGATTAAACAGTTGCAGCATGCCCTGGGAATGCAGATAAAGCGATATCAGACTCA 5967  
QY 82 GACATCTCTGTCCAGTCTCTGGTCCAGATTCACCTGCGCGCCGCCACAGCAGGTGGAG 141  
DB 5968 CAGATCAGCGCTCTCTCCACCATAGCATATATTGGCCACCTGTCTCTCCACAGCC 6027  
QY 142 AGCAGTGACGGGATGGCGCTGGTGGCCCGCAGGCGTGGTGTTCCTCCAAAGAGGAGGAG 201  
DB 6028 CGACTTCACTCTCAGGCGCGGACGAATGCTGGCGACCCCGGCTGAGCAGCGCAGAGGAG 6087  
QY 202 TACTTCAGGTGGATCTAACGACTGTCACCTGGTGGCTCTGGTGGGCAACCCAGGAGCGG 261  
DB 6088 TGGCTGAGGTGACCTCGAAGAGAGCGGTGAAGGTCACAGGCATCACCACCCAGGGCGGTG 6147  
QY 262 CATGCCGGGGCGCTGGCAAGGAGTCTCCCGGAGTACCGGCTGGTACTCCCGGGAT 321  
DB 6148 AAGTCCCTGCTCAGCAGCATGTATGTGAGGAGTTCCTCTGTCTCCAGT---AGTCAGGAC 6204  
QY 322 GGTGCGCGGTGGATGGGTGGAGAGACCGCTGGGGTTCAGAGGTGATCTCAGGCAATGAG 381  
DB 6205 GGCGCGCGGTGGAGCCCTGTGTCTTTCAGGACGGCCACACGAAGGTTTTTTCAGGCAATGAG 6264  
QY 382 GACCTTGAGGAGTGGTGGTGGAGACCTTGGGCCCCCATGTTGCCCGGACTGTTGCG 441

DB 6265 GACTCTCTCCACCCCGTGGTGAACGCTCTGGACCCCGCTGTTCCACGCGCTACTGAGG 6324  
QY 442 TTCTACCCC 450  
DB 6325 ATCCACCCC 6333  
RESULT 9  
US-08-670-707A-38  
Sequence 38, Application US/08670707A  
Patent No. 5859204  
GENERAL INFORMATION:  
APPLICANT: Lollar, John S.  
TITLE OF INVENTION: Hybrid Human/Animal Factor VIII  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Greenlee, Winner and Sullivan, P.C.  
STREET: 5370 Manhattan Circle Suite 201  
CITY: Boulder  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/670,707A  
FILING DATE: 26-JUN-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US94/13200  
FILING DATE: 15-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/212,133  
FILING DATE: 11-MAR-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/864,004  
FILING DATE: 07-APR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Greenlee, Lorraine L.  
REGISTRATION NUMBER: 27,894  
REFERENCE/DOCKET NUMBER: 75-95F  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 303/499-8080  
TELEFAX: 303/499-8089  
INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4334 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: not relevant  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Factor VIII lacking B domain  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 3..4334  
US-08-670-707A-38

Query Match 3.9%; Score 46.6; DB 3; Length 4334;  
Best Local Similarity 45.9%; Pred. No. 0.0059;  
Matches 197; Conservative 0; Mismatches 229; Indels 3; Gaps 1;  
QY 22 TTGTATCTCTCCCAAGTGGCGCTATGCCCTGGCATGCGAGCCGACCATCCAGACAGT 81  
DB 3840 TGTGATTAAACAGTTGCAGCATGCCCTGGGAATGCAGATAAAGCGATATCAGACTCA 3899  
QY 82 GACATCTCTCTCCACCATCTCTGGTCCAGATTCACCTGCGCGCCGCCACAGCAGGTGGAG 141

```
Query Match      3.3%; Score 39.6; DB 4; Length 461;
Best local Similarity 57.1%; Pred. No. 0.17;
Matches 72; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
```

[illegible]

[illegible]

RESULT 12

```

US-08-642-255-48
? Sequence 48, Application US/08642255
? Patent No. 5773249
? GENERAL INFORMATION:
? APPLICANT: CAPPELLO, Joseph
? APPLICANT: FERRARI, Franco A.
? TITLE OF INVENTION: High Molecular Weight Collagen-Like
? TITLE OF INVENTION: Protein Polymers
? NUMBER OF SEQUENCES: 135
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
? STREET: 4 Embarcadero Center, Suite 3400
? CITY: San Francisco
? STATE: California
? COUNTRY: USA
? ZIP: 94111-4187
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentn Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/642,255
? FILING DATE:
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: ROWLAND, Bertram I.
? REGISTRATION NUMBER: 20,015
? REFERENCE/DOCKET NUMBER: A55556-3/BTR
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (415) 494-8700
? TELEFAX: (415) 494-8771
? TELEX: 910 277299 FHT UR
? INFORMATION FOR SEQ ID NO: 48:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 432 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? MOLECULE TYPE: other nucleic acid
? DESCRIPTION: /desc = "synthetic"
US-08-642-255-48

```

[illegible]

Db	189	GGGTGCTCCGGGACCTCGAGCCCGCCAGGTGCGCTTGGACCGCTGGTCCACCGGTGC	24.8
Qy	359	AGGAGGTGATCTCAGGCAATGAGACCCCTGAGGAGTGTGTCTGAAAGGACCTTTGGGCCCC	41.8
Db	249	TCCGGGACCTTCAGGCGCCGACGTGCGCTTGACCGGCTGTGTCTCCACCGGGTCTCCGGG	30.8
Qy	419	CCATGGTTGCCGAGCTGGTTCGGTTCTTACCCCCGGGCTGACCGGGTTCATGAGGCTC	47.4
Db	309	ACCTTCAGGCGCCGACGCTGGCTTGACCGGCTGTGTCTCCACCGGGTGTCTCCGGAC	36.4

RESULT 13

```

US-08-642-255-50
: Sequence 50, Application US/08642255
: Patent No. 5773249
: GENERAL INFORMATION:
: APPLICANT: CAPPELLO, Joseph
: APPLICANT: FERRARI, Franco A.
: TITLE OF INVENTION: High Molecular Weight Collagen-Like
: TITLE OF INVENTION: Protein Polymers
: NUMBER OF SEQUENCES: 135
: CORRESPONDENCE ADDRESS:
: ADDRESS: 4 Embarcadero Center, Suite 3400
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-4187
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC Compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/642,255
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: ROWLAND, Bertram I.
: REGISTRATION NUMBER: 20,015
: REFERENCE/DOCKET NUMBER: A55556-3/BIR
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 494-8700
: TELEFAX: (415) 494-8771
: TELEX: 910 277299 FHT UR
: INFORMATION FOR SEQ ID NO: 50:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 756 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: other nucleic acid
: DESCRIPTION: /desc = "synthetic"
US-08-642-255-50

```

[illegible]





Db 530 TGGTGGCCTGGGGATCTCAGGGAAGGGCTATGGGAGCACGGCGGTCTCCTCAGTGCTGGG 589

Qy 381 GGACCCCTGAGGAGTGTGCTG 402

Db 590 GCTTTCAGGGGCTTGGTACCG 611

Search completed: November 4, 1999, 07:29:41  
Job time: 6329 sec

GenCore version 4.5  
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 4, 1999, 07:32:14 ; Search time 111.15 Seconds  
(without alignments)  
2694.376 Million cell updates/sec

Title: US-08-170-558-7  
Perfect score: 1197  
Sequence: 1 GATGCTGACATGAAGGACA.....CCGAGGGAGCCGACGCGC 1197

Scoring table: IDENTITY\_NUC

Searched: 311585 seqs, 125096042 residues

Database : N\_Geneseq\_36.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1192.2	99.6	3962	1	Q92520	Human mammary carc
2	1192.2	99.6	3962	1	Q92522	Human mammary carc
3	1192.2	99.6	3962	1	T93785	Human mammary carc
4	1171.2	97.8	3754	1	Q84782	Protein-tyrosine-k
5	327.4	27.4	3157	1	Q92521	Human colonic aden
6	327.4	27.4	3157	1	Q92523	Human colonic aden
7	327.4	27.4	3157	1	T93784	CK-2, a human mam
8	327.4	27.4	3096	1	V48282	Discoilin domain r
9	321	26.8	3120	1	V55895	Receptor protein t
10	321	26.8	3120	1	V65317	Receptor protein t
11	143.2	12.0	2128	1	Q64158	Partial coding seq
12	73.8	6.2	4451	1	V23288	Synthetic human Fa
13	46.6	3.9	6402	1	V12113	Sus scrofa factor
14	45.6	3.9	4334	1	V12112	Homo sapiens facto
15	43.2	3.6	2730	1	X09001	Human semaphorin r
16	43.2	3.6	2781	1	X09002	Human semaphorin r
17	41	3.4	549	1	X08979	Human monocyte gro
18	40	3.3	3539	1	X08996	Rat semaphorin rec
19	40	3.3	4718	1	X08987	Mouse semaphorin r
20	40	3.3	4733	1	X08988	Mouse semaphorin r
21	40	3.3	4769	1	X08989	Mouse semaphorin r
22	40	3.3	4784	1	X09000	Mouse semaphorin r
23	40	3.3	4765	1	X09003	Mouse semaphorin r
24	40	3.3	4780	1	X09004	Mouse semaphorin r
25	39.6	3.3	461	1	X25867	Clone P8Gen-4, hom
26	37.6	3.1	432	1	Q43031	Collagen-like poly
27	37.6	3.1	756	1	Q43032	Collagen-like poly
28	37.6	3.1	432	1	T16765	Collagen-like poly
29	37.6	3.1	756	1	T16766	Collagen-like poly
30	37.2	3.1	1116	1	T61096	N-acetylglucosamin
31	37.2	3.1	114955	1	X53491	Human adenosine A1
32	37	3.1	114955	1	X53491	Human adenosine A1
33	36.6	3.1	3728	1	X58708	Mouse OSF-5 coding
34	36.6	3.1	3633	1	T97611	Mouse E2A-binding
35	36.2	3.0	10596	1	Q51731	Plasmid pCisEBON f
36	36.2	3.0	10596	1	T40348	Plasmid pCisEBON f
37	36.2	3.0	9600	1	V21683	Vector plasmid pCM
38	36.2	3.0	10596	1	X15650	Nucleotide sequenc
39	36	3.0	2832	1	Q50631	Brain factor-1. Is
40	36	3.0	799	1	V55831	Nucleotide sequenc
41	35.8	3.0	2103	1	V18444	Human placental-de
42	35.8	3.0	3402	1	V93909	Human bone and pro
43	35.6	3.0	1817	1	Q50511	Bx1 gene. DNA sequ

44 35.4 3.0 2846 1 V41260 Chlamydomonas rein  
45 ...35.4 3.0 1278 1 V41263 Chlamydomonas rein

## ALIGNMENTS

### RESULT 1

Q92520 ID Q92520 standard; cDNA to mRNA; 3962 BP.  
AC Q92520;  
DE 26-NOV-1995 (first entry)  
DE Human mammary carcinoma kinase 10 (MCK-10) cDNA.  
KW Mammary carcinoma kinase 10; transmembrane receptor;  
KW receptor tyrosine kinase; cancer; ss.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT cds 321..3080  
FT misc\_difference 2315  
FT /\*tag= a  
FT /\*tag= b  
FT /\*note= "some clones have 6 AA deletion here"  
PN W09514088-A.  
PD 26-MAY-1995.  
PF 16-NOV-1994; E03797.  
PR 16-NOV-1993; US-153397.  
PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.  
PI Alves FHE, Ullrich A;  
DR WPI; 95-224054/29.  
DR P-PSDB; R73502.  
PT New nucleic acid encoding MCK-10 receptor tyrosine kinase - and  
PT derived vectors, transformed cells, proteins and antibodies useful  
PT for diagnosis and treatment of proliferative disease, esp. cancer,  
PT and for screening modulators  
PS Claim 4; Page 50-52; 115pp; English.  
CC cDNA prep'd. from human breast cancer cell line MCF7 (ATCC HTB22) and  
CC used in a PCR with two degenerate oligo primer pools based on  
CC conserved sequences of the kinase domain of receptor tyrosine  
CC kinases. One clone, designated MCK-10, was identified as novel RTK.  
CC The PCR fragment was used to screen a lambda gill library of human  
CC fetal brain cDNA. Several overlapping clones were identified. The  
CC composite of these cDNA clones is given in Q92520 and the deduced AA  
CC sequence in R73502. Some of the clones had a deletion of 6 AAs at  
CC posn. 2315 in the MCK-10 sequence. MCK-10 has all the  
CC characteristics of a receptor PTK (see R73502 FT). Screening of  
CC human placental library yielded two cDNA clones MCK-10-1 and  
CC MCK-10-2. One of the clones isolated from the human fetal brain  
CC library contd. an additional 18 nts in the TK domain. The MCK-10 splice  
CC isoforms have been designated MCK-10-1 (with an additional 111 bp between  
CC nts 1832 and 1943); MCK-10-2 (without any insertions); MCK-10-3 (with the  
CC additional 111 bp and and 18 bp in the TK domain); and MCK-10-4 (with the  
CC additional 18 bp). The predicted mol. wts. of MCK-10-1 and MCK-10-2  
CC proreceptors are 101.13 and 97.17 kD respectively, and can thus be  
CC subdivided into a 34.31 kD alpha subunit and and 66.84 or 62.88 kD  
CC beta subunits that contain the TK homology and alternative splice sites.  
SQ Sequence 3962 BP; 735 A; 1235 C; 1181 G; 811 T;

Query Match 99.6%; Score 1192.2; DB 1; Length 3962;  
Best Local Similarity 99.7%; Pred. No. 6.4e-274;  
Matches 1194; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GATGCTGACATGAAGGACATTTTATCTCTGCAAGTCGCGATGCGCTGGCGATGCAG 60  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
375 GATGCTGACATGAAGGACATTTTATCTCTGCAAGTCGCGATGCGCTGGCGATGCAG 434  
QY 61 GATCGGACCATCCAGACAGTGCATCTCTGCTCCAGTCTCTGCTGATGATCCACTGCC 120  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
435 GATCGGACCATCCAGACAGTGCATCTCTGCTCCAGTCTCTGCTGATGATCCACTGCC 494  
QY 121 GCCCGCCACAGAGGTTGGAGAGCAGTGCAGGGGATGGGGCTGTGTGCCCGCAGGGTGC 180  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
495 GCCCGCCACAGAGGTTGGAGAGCAGTGCAGGGGATGGGGCTGTGTGCCCGCAGGGTGC 554

QY 181 GTGTTTCCCAAGGAGGAGTACTTGCAGTGGATCTACAAGCACTGCACCTGGTGGCT 240  
DB 555 GTGTTTCCCAAGGAGGAGTACTTGCAGTGGATCTACAAGCACTGCACCTGGTGGCT 614  
QY 241 CTGTGGGACCCAGGACGAGCATGCCGGGGGCTGGGCAAGAGTTCTCCCGAGCTAC 300  
DB 615 CTGTGGGACCCAGGACGAGCATGCCGGGGGCTGGGCAAGAGTTCTCCCGAGCTAC 674  
QY 301 CGGCTGGTTACTCCCGGATGTCGCGCTGGATGGCTGGAGGACCGCTGGGTGTCAG 360  
DB 675 CGGCTGGTTACTCCCGGATGTCGCGCTGGATGGCTGGAGGACCGCTGGGTGTCAG 734  
QY 361 GAGGTGATCTCAGGCAATAGGAGCCCTGAGGGAGTGTGCTGAAGGACCTTGGGCCCCC 420  
DB 735 GAGGTGATCTCAGGCAATAGGAGCCCTGAGGGAGTGTGCTGAAGGACCTTGGGCCCCC 794  
QY 421 ATGTTGCCGCACTGGTTCCTTACCCCGGGCTGACCGGGTCAATGAGTGTCTGTG 480  
DB 795 ATGTTGCCGCACTGGTTCCTTACCCCGGGCTGACCGGGTCAATGAGTGTCTGTG 854  
QY 481 CGGCTAGAGCTCTATGCTGCTTGGAGGATGAGTCTGCTTACACGCGCCCTGTG 540  
DB 855 CGGCTAGAGCTCTATGCTGCTTGGAGGATGAGTCTGCTTACACGCGCCCTGTG 914  
QY 541 GGGCAGACAATGTTATCTGAGCGCGTGTACCTCAACGACTCCACCTATGACGGACAT 600  
DB 915 GGGCAGACAATGTTATCTGAGCGCGTGTACCTCAACGACTCCACCTATGACGGACAT 974  
QY 601 ACCGTGGCGGACTGCAATATGGGGTCTGGGCCAGCTGGCAGATGTTGTGGGGCTG 660  
DB 975 ACCGTGGCGGACTGCAATATGGGGTCTGGGCCAGCTGGCAGATGTTGTGGGGCTG 1034  
QY 661 GATGACTTTAGAGAGTCAGAGCTCGGGTCTGGCCAGCTATGACTATGTTGGGATGG 720  
DB 1035 GATGACTTTAGAGAGTCAGAGCTCGGGTCTGGCCAGCTATGACTATGTTGGGATGG 1094  
QY 721 AGCAACACAGCTCTCCAGTGGTATGTGGAGATGGATTTGAGTTTGCAGCGCTGAGG 780  
DB 1095 AGCAACACAGCTCTCCAGTGGTATGTGGAGATGGATTTGAGTTTGCAGCGCTGAGG 1154  
QY 781 GCTTCCAGGCTATGCAAGTCACTGTAAACAATGACACAGCTGGGAGCGCTGCTGCT 840  
DB 1155 GCTTCCAGGCTATGCAAGTCACTGTAAACAATGACACAGCTGGGAGCGCTGCTGCT 1214  
QY 841 GCGGGGTGGAATGTCGCTTCCGGCTGGCCCTGCCATGGCTGGGAGGGAGCCCATG 900  
DB 1215 GCGGGGTGGAATGTCGCTTCCGGCTGGCCCTGCCATGGCTGGGAGGGAGCCCATG 1274  
QY 901 CGCCACAACCTAGGGGGCAACCTGGGGGACCCAGAGCCCGGGTGTCTCAGTGCCTTT 960  
DB 1275 CGCCACAACCTAGGGGGCAACCTGGGGGACCCAGAGCCCGGGTGTCTCAGTGCCTTT 1334  
QY 961 GCGGCGCTGTGCTGCTTTCAGTGGCGCTTCTTTCGCGGGGCTGCTGTTACTC 1020  
DB 1335 GCGGCGCTGTGCTGCTTTCAGTGGCGCTTCTTTCGCGGGGCTGCTGTTACTC 1394  
QY 1021 TTCAGCAAAATCTCCTTCTATCTGTGTGAACAATTCCTCTCGGGCACTGGGAGGC 1080  
DB 1395 TTCAGCAAAATCTCCTTCTATCTGTGTGAACAATTCCTCTCGGGCACTGGGAGGC 1454  
QY 1081 ACCTTCCGCGCAAGCCCTGTGGCGGCTGGCCCACTCCACCAACTTCAGCAGCTTG 1140  
DB 1455 ACCTTCCGCGCAAGCCCTGTGGCGGCTGGCCCACTCCACCAACTTCAGCAGCTTG 1514  
QY 1141 GAGCTGGAGCCCAAGAGGACGAGCCCTGGCCAAAGCCCGAGGGAGCCGACCGCC 1197  
DB 1515 GAGCTGGAGCCCAAGAGGACGAGCCCTGGCCAAAGCCCGAGGGAGCCGACCGCC 1571

RESULT 2

Q2522

ID Q2522 standard; cDNA to mRNA; 3962 BP.

Q2522;  
DT 26-NOV-1995 (first entry)  
DE Human mammary carcinoma kinase 10 (MCK-10) cDNA.  
KW Mammary carcinoma kinase 10; MCK-10; transmembrane receptor;  
KW receptor tyrosine kinase; cancer; ss.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT 321..3080  
FT misc\_difference 2315  
FT /\*tag= a  
FT /\*tag= b  
FT /\*note= "some clones have AA deletion here"  
PN W09514089-A.  
PD 26-MAY-1995.  
PF 16-NOV-1994; E03799.  
PR 16-NOV-1993; US-153397.  
PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.  
PI Alves FHE, Ullrich A;  
DR WPI: 95-224055/29.  
DR P-PSDB; R75504.  
PT New nucleic acid encoding CCK-2 receptor tyrosine kinase - and  
PT derived vectors, transformed cells, proteins and antibodies, useful  
PT for diagnosis and treatment of proliferative and nervous system  
PT diseases and for screening modulators.  
PS Disclosure; Page 67-69; 115pp; English.  
CC cDNA prep. from human breast cancer cell line MCF7 (ATCC HTB22) was  
CC used in a PCR with two degenerate oligo primer pools based on  
CC conserved sequences of the kinase domain of receptor tyrosine  
CC kinases. One clone, designated MCK-10, was identified as novel RTK.  
CC The PCR fragment was used to screen a lambda gtl1 library of human  
CC fetal brain cDNA. Several overlapping clones were identified. The  
CC composite of these cDNA clones is given in Q2522 and the deduced AA  
CC sequence in R75504. Some of the clones had a deletion of 6AA at posn.  
CC 2315 in the MCK-10 sequence. MCK-10 has all the characteristics of  
CC a receptor PTK (see R75504 FT). Screening of human placental library  
CC yielded two cDNA clones. One of the clones isolated from the human  
CC fetal brain library contained an additional 18 nts in the TK  
CC domain. The MCK-10 splice isoforms have been designated MCK-10-1  
CC (with an additional 111 bp between nts 1832 and 1943); MCK-10-2  
CC (without any insertions); MCK-10-3 (with the additional 111 bps and  
CC 18 bp in the TK domain); and MCK-10-4 (with the additional 18 bp).  
CC The predicted mol. wts. of MCK-10-1 and MCK-10-2 precursors are  
CC 101.13 and 97.17 kD respectively, and can thus be subdivided into a  
CC 34.31 kD alpha subunit and a 66.84 or 62.88 kD beta subunits that  
CC contain the TK homology and alternative splice sites.  
SQ Sequence 3962 BP; 735 A; 1235 C; 1181 G; 811 T;

Query Match 99.6%; Score 1192.2; DB 1; Length 3962;

Best Local Similarity 99.7%; Pred. No. 6.4e-274;

Matches 1194; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GATGCTGACATCAAGGACATTTTGCATCTGCCAAGTCCCGTATGCCCTGGGCATGCAG 60  
DB 375 GATGCTGACATCAAGGACATTTTGCATCTGCCAAGTCCCGTATGCCCTGGGCATGCAG 434  
QY 61 GACCGGACCATCCAGACAGTGCATCTCTGCTTCCAGCTCCTGCTCAGATTCCACTGCC 120  
DB 435 GACCGGACCATCCAGACAGTGCATCTCTGCTTCCAGCTCCTGCTCAGATTCCACTGCC 494  
QY 121 GCGCGCCACAGAGGTTGGAGAGCAGTGCAGGGGATGGGGCTGTGCCCCCAGGGTGC 180  
DB 495 GCGCGCCACAGAGGTTGGAGAGCAGTGCAGGGGATGGGGCTGTGCCCCCAGGGTGC 554  
QY 181 GTGTTTCCCAAGGAGGAGTACTTGCAGTGGATCTACAACGACTGCACCTGGTGGCT 240  
DB 555 GTGTTTCCCAAGGAGGAGTACTTGCAGTGGATCTACAACGACTGCACCTGGTGGCT 614  
QY 241 CTGTGGGACCCAGGACGAGCATGCCGGGGGCTTGGGCAAGAGTTCTCCCGAGCTAC 300  
DB 615 CTGTGGGACCCAGGACGAGCATGCCGGGGGCTTGGGCAAGAGTTCTCCCGAGCTAC 674  
QY 301 CGGCTGGTTACTCCCGGATGTCGCGCTGGATGGCTGGAGGACCGCTGGGTGTCAG 360

```

Db 675 CGGCTCGTACTCCCGGATGGTCCCGCTGGATGGGCTGAAGAGCCGCTGGGGTCTAG 734
QY 361 GAGGTGATCTCAGGAATAGGACCTTGAAGAGTGGTGGTGAAGAGACCTTTGGGCCCCC 420
Db 735 GAGGTGATCTCAGGAATAGGACCTTGAAGAGTGGTGGTGAAGAGACCTTTGGGCCCCC 794
QY 421 ATGGTTGCCCGACTGGTTCGCTTACCCCGGCTGACCGGGTCTAGGAGTCTGTCTG 480
Db 795 ATGGTTGCCCGACTGGTTCGCTTACCCCGGCTGACCGGGTCTAGGAGTCTGTCTG 854
QY 481 CGGTAGAGCTCTATGGTTCCTTGGAGGATGGACTCTTCTTACACCGCCCTGTG 540
Db 855 CGGTAGAGCTCTATGGTTCCTTGGAGGATGGACTCTTCTTACACCGCCCTGTG 914
QY 541 GGGCAGACAATGATTTATCTAGCGCGGTGTACCTCAACGACTCCACCTATGACGGACAT 600
Db 915 GGGCAGACAATGATTTATCTAGCGCGGTGTACCTCAACGACTCCACCTATGACGGACAT 974
QY 601 ACCGTGGGGGACTCAGTATGGGGTCTGGCCAGCTGGCCAGCTATGACTATGTGGGATGG 560
Db 975 ACCGTGGGGGACTCAGTATGGGGTCTGGCCAGCTGGCCAGCTATGACTATGTGGGATGG 1034
QY 661 GATGACTTTAGGAAGAGTCAAGAGTGGGGTCTGGCCAGCTATGACTATGTGGGATGG 720
Db 1035 GATGACTTTAGGAAGAGTCAAGAGTGGGGTCTGGCCAGCTATGACTATGTGGGATGG 1094
QY 721 ACACACACAGCTTCTCAGTGGCTATGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 780
Db 1095 ACACACACAGCTTCTCAGTGGCTATGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 1154
QY 781 GCTTCCAGGCTATCAGGTCACCTGTAAACATGACACAGCTGGGAGCCGCTGTGCT 840
Db 1155 GCTTCCAGGCTATCAGGTCACCTGTAAACATGACACAGCTGGGAGCCGCTGTGCT 1214
QY 841 GCGGGGTGGAATGTCGCTTCGCGGCTGGCCCTGCGATGGCTGGGAGGGAGCCCATG 900
Db 1215 GCGGGGTGGAATGTCGCTTCGCGGCTGGCCCTGCGATGGCTGGGAGGGAGCCCATG 1274
QY 901 GCGCACACCTAGGGGGCAACCTGGGGGAGCCAGAGCCCGGGGTGTCTCAGTCCGCCCTT 960
Db 1275 GCGCACACCTAGGGGGCAACCTGGGGGAGCCAGAGCCCGGGGTGTCTCAGTCCGCCCTT 1334
QY 961 GCGCGCGCTGTGGCTGCTTCTGAGTGGCGCTTCTCTTGGGGGCGCTGTGTTACTC 1020
Db 1335 GCGCGCGCTGTGGCTGCTTCTGAGTGGCGCTTCTCTTGGGGGCGCTGTGTTACTC 1394
QY 1021 TTCAGCAAAATCTCCTTCTATCTGATGTGTGAACAAATCTCTCCGGCACTGGGAGGC 1080
Db 1395 TTCAGCAAAATCTCCTTCTATCTGATGTGTGAACAAATCTCTCCGGCACTGGGAGGC 1454
QY 1081 ACCTTCCCGCAGCCCTGTGGCGCTGGCCGCTGAGCCACCTCCACAACTTCAGCAGCTG 1140
Db 1455 ACCTTCCCGCAGCCCTGTGGCGCTGGCCGCTGAGCCACCTCCACAACTTCAGCAGCTG 1514
QY 1141 GAGCTGGAGCCAGAGGCGCAGCCGCTGGCCGCTGGCCGCTGGCCGCTGGCCGCTGGCC 1197
Db 1515 GAGCTGGAGCCAGAGGCGCAGCCGCTGGCCGCTGGCCGCTGGCCGCTGGCCGCTGGCC 1571

```

RESULT 3

T93785  
ID T93785 standard; cDNA; 3962 BP.  
AC T93785.  
DE 16-FEB-1998 (first entry)  
DE Human mammary carcinoma kinase 10 (MCK-10) cDNA sequence.  
KW Mammary carcinoma kinase; MCK-10; receptor tyrosine kinase;  
KW proliferative disease; cancer; insulin receptor family;  
KW tyrosine kinase neurotrophin receptor; MCK-10 activity;  
KW neurological disorder; aberrant expression; ds.  
OS Homo sapiens.  
FH Key  
FT Location/Qualifiers  
321..3077

FT PN /\*tag= a

US5677144-A.  
14-OCT-1997.  
08-NOV-1994; 336343.  
16-NOV-1993; US-153397.  
(ALVE/) ALVES F H E.  
PA (ULLR/) ULLRICH A.  
PI ALVES FHE, Ullrich A;  
DR WPI; 97-511869/47.  
DR P-PSDB; W34672.

Truncated receptor tyrosine kinase CCK-2 - and nucleic acid coding  
for it, useful for cancer diagnosis  
PT Disclosure; Fig 1; 70pp; English.

CC The present sequence represents the cDNA of a mammary carcinoma kinase,  
CC called MCK-10. This kinase belongs to a novel family of receptor tyrosine  
CC kinases, and expression is associated with proliferative diseases such as  
CC cancer. The MCK-10 receptor tyrosine kinase has extensive sequence  
CC similarity to the insulin receptor family. The MCK-10 gene was isolated  
CC by PCR using 2 degenerate oligonucleotide primer pools, using a template  
CC cDNA synthesised by reverse transcription of poly-A RNA from the human  
CC mammary carcinoma cell line MCF7. MCK-10 is expressed in brain tissue,  
CC and the protein shares homology with the tyrosine kinase neurotrophin  
CC receptor. Modulation of MCK-10 activity therefore may be used for  
CC treatment of neurological disorders. MCK-10 is also expressed in a  
CC variety of cancer cell lines and tumour tissue. The present sequence,  
CC or parts of it, can be used for diagnostic purposes to detect aberrant  
CC expression of MCK-10 genes. Inhibitors of MCK-10 receptor activity may  
CC have therapeutic value in the treatment of diseases such as cancer.  
SQ Sequence 3962 BP; 735 A; 1234 C; 1182 G; 811 T;

Query Match

Best Local Similarity 99.6%; Score 1192.2; DB 1; Length 3962;  
Matches 1194; Conservative 99.7%; Pred. No. 6.4e-274;

Mismatches 0; Indels 0; Gaps 0;

QY	1	GATGCTGACATGAGGACATTTTGTATCTGCTCCAGTGCCTGATGCCCTGGCGATGAC	60
Db	375	GATGCTGACATGAGGACATTTTGTATCTGCTCCAGTGCCTGATGCCCTGGCGATGAC	434
QY	61	GACCGGACCATCCAGACAGTGCATCTCTGCTCCAGTGCCTGATGCCCTGGCGATGAC	120
Db	435	GACCGGACCATCCAGACAGTGCATCTCTGCTCCAGTGCCTGATGCCCTGGCGATGAC	494
QY	121	GCCCGCCACAGAGTGGAGAGCAGTGCAGGGGATGGGGCTGGTGGCCCGCAGAGGTG	180
Db	495	GCCCGCCACAGAGTGGAGAGCAGTGCAGGGGATGGGGCTGGTGGCCCGCAGAGGTG	554
QY	181	GTGTTTCCCAAGGAGGAGGAGTACTTGCAGTGGATCTACACGACTGCACCTGGTGGCT	240
Db	555	GTGTTTCCCAAGGAGGAGGAGTACTTGCAGTGGATCTACACGACTGCACCTGGTGGCT	614
QY	241	CTGGTGGCCACCCAGGAGCGCATGCCGGGCGCTGGGCAAGGAGTTCCTCCCGAGCTAC	300
Db	615	CTGGTGGCCACCCAGGAGCGCATGCCGGGCGCTGGGCAAGGAGTTCCTCCCGAGCTAC	674
QY	301	CGGCTGCGTTACTCCCGGATGCTGCCGCTGGATGGGCTGGAAGGAGCGCTGGGGTCTAG	360
Db	675	CGGCTGCGTTACTCCCGGATGCTGCCGCTGGATGGGCTGGAAGGAGCGCTGGGGTCTAG	734
QY	361	GAGGTGATCTCAGGCAATGAGGACCTGAGGAGTGGTGTGTAAGGACCTTGGGCCCCC	420
Db	735	GAGGTGATCTCAGGCAATGAGGACCTGAGGAGTGGTGTGTAAGGACCTTGGGCCCCC	794
QY	421	ATGGTTGCCCGACTGGTTCGCTTCTACCCCGGCTGACCGGGTCTAGGAGTCTGTCTG	480
Db	795	ATGGTTGCCCGACTGGTTCGCTTCTACCCCGGCTGACCGGGTCTAGGAGTCTGTCTG	854
QY	481	CGGGTAGAGCTCTATGGTTCCTCTGGAGGATGAGCTCTCTTTACACCGCCCTGTG	540
Db	855	CGGGTAGAGCTCTATGGTTCCTCTGGAGGATGAGCTCTCTTTACACCGCCCTGTG	914
QY	541	GGGCAGACAATGATTTATCTAGCGCGGTGTACCTCAACGACTCCACCTATGACGGACAT	600
Db		GGGCAGACAATGATTTATCTAGCGCGGTGTACCTCAACGACTCCACCTATGACGGACAT	

Db 915 GGCAGAGAAATGATTTATCTAGGCGCTGTACCTCAAGACTCCACCTATGACGGACAT 974  
 Qy 601 ACCGTGGCGGAGCTGACAGTATGGGGTCTGGCGAGCTGCAGATGTTGGTGGGGCTG 660  
 Db 975 ACCGTGGCGGAGCTGACAGTATGGGGTCTGGCGAGCTGCAGATGTTGGTGGGGCTG 1034  
 Qy 661 GATGACTTTAGGAGAGTACAGGAGCTGCGGGTCTGGCCAGCTATGACATGTTGGGATG 720  
 Db 1035 GATGACTTTAGGAGAGTACAGGAGCTGCGGGTCTGGCCAGCTATGACATGTTGGGATG 1094  
 Qy 721 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGAGTTTGTAGTTTACCGGCTGAGG 780  
 Db 1095 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGAGTTTGTAGTTTACCGGCTGAGG 1154  
 Qy 781 GCGTTCAGGCTATGACAGTCCACTGTAAACAATGACACAGCTGGGAGCGGCTGTGCT 840  
 Db 1155 GCGTTCAGGCTATGACAGTCCACTGTAAACAATGACACAGCTGGGAGCGGCTGTGCT 1214  
 Qy 841 GCGGGGTGGAATGCTCTCCGCGTGGCCCTGCCATGSCCTGGAGGGGAGCCCATG 900  
 Db 1215 GCGGGGTGGAATGCTCTCCGCGTGGCCCTGCCATGSCCTGGAGGGGAGCCCATG 1274  
 Qy 901 CGGCACAACTAGGGGAGCACTGGGGGAGCCAGAGCCCGGCTGTCTCAGTGCCCTT 960  
 Db 1275 CGGCACAACTAGGGGAGCACTGGGGGAGCCAGAGCCCGGCTGTCTCAGTGCCCTT 1334  
 Qy 961 GCGGCGGTGGCTGCTCTTCTGACGTGCGCTTCTTTCGGGGGCGCTGTGTTACTC 1020  
 Db 1335 GCGGCGGTGGCTGCTCTTCTGACGTGCGCTTCTTTCGGGGGCGCTGTGTTACTC 1394  
 Qy 1021 TTCAGGAAATCTCTTCTATCTCTGATGTGGTGAACAATTCCTCTCGGCACTGGGAGGC 1080  
 Db 1395 TTCAGGAAATCTCTTCTATCTCTGATGTGGTGAACAATTCCTCTCGGCACTGGGAGGC 1454  
 Qy 1081 ACCTTCGCGGAGCCCTGTGTGGCGGCTGGCCAGCTCCACCACTTCACAGCTTG 1140  
 Db 1455 ACCTTCGCGGAGCCCTGTGTGGCGGCTGGCCAGCTCCACCACTTCACAGCTTG 1514  
 Qy 1141 GAGCTGAGCCAGGAGCCAGCAGCCGCTGGCCAGCCGAGGGAGCCGAGCCGCGC 1197  
 Db 1515 GAGCTGAGCCAGGAGCCAGCAGCCGCTGGCCAGCCGAGGGAGCCGAGCCGCGC 1571

RESULT 4

ID Q84782  
 AC Q84782 standard; DNA; 3754 BP.  
 DT 17-AUG-1995 (first entry)  
 DE Protein-tyrosine-kinase PTK22.  
 KW Protein-tyrosine-kinase; PTK; discolidin domain receptor; cancer;  
 KW breast tumor; mammary carcinoma; diagnosis; prognosis; therapy; ds.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT cds 142..2886  
 FT /\*tag= a  
 PN W09502187-A.  
 PD 19-JAN-1995.  
 PF 08-JUL-1994; G01480.  
 PR 09-JUL-1993; GB-014271.  
 PA (CANC-) CANCER RES INST.  
 PA (WELL) WELLCOME FOUND LTD.  
 PI Barker KT, Crompton MR, Gusterson BA, Martindale JE;  
 PI Mitchell PJ, Page M, Spence P;  
 DR WPI: 95-066991/09.  
 DR P-PSDB: R71100.  
 PT Method for screening substances, using protein tyrosine kinase -  
 PT for potential utility as therapeutic agents for cancer  
 PS Claim 1; Page 26-30; 51pp; English.  
 CC cDNA derived from tumor metastatic tissue was amplified using  
 CC primers (given in Q84783-84) based on sequences (R71101, R71103)  
 CC associated with protein-tyrosine-kinases (PTK). Novel PTK22 was  
 CC identified in an isolated subclone. The 3' sequence of PTK22 was  
 CC obtained by reverse transcription (using the primer of Q84786) and

CC PCR amplification (primers Q84787-88) of RNA of human breast  
 CC carcinoma cell line MDA MG 468. The partial DNA sequence of PTK22  
 CC is given in Q84782.  
 SQ Sequence 3754 BP; 713 A; 1145 C; 1121 G; 775 T;  
 Query Match 97.8%; Score 1171.2; DB 1; Length 3754;  
 Best Local Similarity 99.1%; Pred. No. 6e-269;  
 Matches 1189; Conservative 0; Mismatches 8; Indels 3; Gaps 1;  
 Qy 1 GATGCTGACATGAAGGACATTTTATCTCTCCAGTGGCCGCTATGCTGGGATCGAG 60  
 Db 196 GATGCTGACATGAAGGACATTTTATCTCTCCAGTGGCCGCTATGCTGGGATCGAG 255  
 Qy 61 GACCGGACATCCAGACAGTACATCTCTCTCCAGTGGCCGCTATGCTGGGATCGAG 120  
 Db 256 GACCGGACATCCAGACAGTACATCTCTCTCCAGTGGCCGCTATGCTGGGATCGAG 315  
 Qy 121 GCGCGCCACACAGAGTTGGAGAGCAGTACGCGGATGGGCTGTGGCCGCGGAGGTCG 180  
 Db 316 GCGCGCCACACAGAGTTGGAGAGCAGTACGCGGATGGGCTGTGGCCGCGGAGGTCG 375  
 Qy 181 GTGTTTCCAGAGGAGGAGTACTTGCAGTGGATCTACAACTGACCTGCTGGTGGCT 240  
 Db 376 GTGTTTCCAGAGGAGGAGTACTTGCAGTGGATCTACAACTGACCTGCTGGTGGCT 435  
 Qy 241 CTGTTGGGACCCAGGAGCGCATGCGGGGCGCTGGGCAAGAGTCTTCCCGGAGGTAC 300  
 Db 436 CTGTTGGGACCCAGGAGCGCATGCGGGGCGCTGGGCAAGAGTCTTCCCGGAGGTAC 495  
 Qy 301 CGGCTGCGTTACTCCCGGATGGTGGCGCTGGATGGGCTGGAAGACCGCTGGGCTGAG 360  
 Db 496 CGGCTGCGTTACTCCCGGATGGTGGCGCTGGATGGGCTGGAAGACCGCTGGGCTGAG 555  
 Qy 361 GAGGTGATCTCAGCAATGAGGACCTCAGGAGTGGTGGCTGGAAGACCTTGGGCCCCC 420  
 Db 556 GAGGTGATCTCAGCAATGAGGACCTCAGGAGTGGTGGCTGGAAGACCTTGGGCCCCC 615  
 Qy 421 ATGTTGGCGGAGTCTGCTTCTTACCCCGGCGTACCGGGCTCATGAGGCTCTGTCTG 480  
 Db 616 ATGTTGGCGGAGTCTGCTTCTTACCCCGGCGTACCGGGCTCATGAGGCTCTGTCTG 675  
 Qy 481 CGGGTAGAGCTCTATGGCTGCTCTGGAGGATGAGTCTCTTTACACCGCCCTGTG 540  
 Db 676 CGGGTAGAGCTCTATGGCTGCTCTGGAGGATGAGTCTCTTTACACCGCCCTGTG 735  
 Qy 541 GGGCAGACAATGATTTATCTGAGGCGCTGTACCTCAACGACTCCACCTATGACGGACAT 600  
 Db 736 GGGCAGACAATGATTTATCTGAGGCGCTGTACCTCAACGACTCCACCTATGACGGACAT 795  
 Qy 601 ACCGTGGCGGAGTCCAGTATGGGGTCTGGGCCAGCTGGCAGATGTTGGTGGGGCTG 660  
 Db 796 ACCGTGGCGGAGTCCAGTATGGGGTCTGGGCCAGCTGGCAGATGTTGGTGGGGCTG 855  
 Qy 661 GATGACTTTAGGAAGAGTACAGGAGTGGGGTCTGGCCAGGCTATGACTATGTTGGATG 720  
 Db 856 GATGACTTTAGGAAGAGTACAGGAGTGGGGTCTGGCCAGGCTATGACTATGTTGGATG 915  
 Qy 721 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGAGTTTGTAGTTTACCGGCTGAGG 780  
 Db 916 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGAGTTTGTAGTTTACCGGCTGAGG 975  
 Qy 781 GCGTTCAGGCTATGACAGTCCACTGTAAACAATGACACAGCTGGGAGCGGCTGTGCT 840  
 Db 976 GCGTTCAGGCTATGACAGTCCACTGTAAACAATGACACAGCTGGGAGCGGCTGTGCT 1035  
 Qy 841 GCGGGGTGGAATGCTCTCCGCGTGGCCCTGCCATGSCCTGGAGGGGAGCCCATG 900  
 Db 1036 GCGGGGTGGAATGCTCTCCGCGTGGCCCTGCCATGSCCTGGAGGGGAGCCCATG 1095  
 Qy 901 CGGCACAACTAGGGGAGCACTGGGGGAGCCAGAGCCCGGCTGTCTCAGTGCCCTT 960  
 Db 1096 CGGCACAACTAGGGGAGCACTGGGGGAGCCAGAGCCCGGCTGTCTCAGTGCCCTT 1155

```

QY 961 GCGGCGCGTGGTGGCTTTCTGAGTGGCGCTTCTTTTGGGGGCGCTGTTACTC 1020
Db 1156 GCGGCGCGTGGTGGCTTTCTGAGTGGCGCTTCTTTTGGGGGCGCTGTTACTC 1215
QY 1021 TTCACGGAATCTCTTCATCTCTGATGTGTGAACAAATTCCTCCGGCACTGGAGGC 1080
Db 1216 TTCACGGAATCTCTTCATCTCTGATGTGTGAACAAATTCCTCCGGCACTGGAGGC 1275
QY 1081 ACCTTCCCGCCAGCCCTTGGTGGCGGCTGGCCACCTCCACCAATTCAGCAGCTTG 1140
Db 1276 ACCTTCCCGCCAGCCCTTGGTGGCGGCTGGCCACCTCCACCAATTCAGCAGCTTG 1335
QY 1141 GAGCTGGAGCCAGAGCCAGCAGC---AGCCGTGGCCAAAGCCCGAGGGAGCCCGACCGCC 1197
Db 1336 GAGCTGGAGCCAGAGCCAGCAGCAGCAGCCGCTGGCCAAAGCCCGAGGGAGCCCGACCGCC 1395

RESULT 5
Q92521
ID Q92521 standard; cDNA to mRNA; 3157 BP.
AC Q92521;
DT 26-NOV-1995 (first entry)
DE Human colonic adenocarcinoma kinase 2 (CKK-2) cDNA.
KW Mammary carcinoma kinase 10; MCK-10; transmembrane receptor; CKK-2;
OS Homo sapiens.
FH Key
FT Location/Qualifiers
FT cds
FT 370..2937
FT /*tag= a
PN W09514088-A.
PD 26-MAY-1995.
PE 16-NOV-1994; E03797.
PR 16-NOV-1993; US-153397.
PA (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
PI Alves PHE Ullrich A;
DR WPI: 95-224054/29.
DR P-PSDB; R75503.
PT New nucleic acid encoding MCK-10 receptor tyrosine kinase - and
PT derived vectors, transformed cells, proteins and antibodies useful
PT for diagnosis and treatment of proliferative disease, esp. cancer,
PT and for screening modulators
PS Disclosure: Page 56-57; 11pp; English.
CC A member of the mammary carcinoma kinase 10 (MCK-10) receptor
CC tyrosine kinase family was identified using a PCR (with two degenerate
CC oligo primer pools based on conserved sequences of the kinase domain of
CC receptor tyrosine kinases) and cDNA prep. from colonic
CC adenocarcinoma RNA. The nt sequence of the novel receptor,
CC designated CKK-2, is given in Q92521 and the deduced AA sequence in
CC R75503. Analysis of the CKK-2 nt and AA sequence indicated
CC significant homology with MCK-10 throughout the extracellular,
CC transmembrane and intracellular regions. The regions of homology
CC extend into the N-terminus consensus sequence for the discoidin I
CC like family of proteins.
SQ Sequence 3157 BP; 790 A; 802 C; 759 G; 806 T;

Query Match 27.4%; Score 327.4; DB 1; Length 3157;
Best Local Similarity 58.8%; Pred. No. 6.7e-69;
Matches 644; Conservative 0; Mismatches 436; Indels 15; Gaps 4;

QY 13 AAGGGACATTTGATCTTCCAGTGCCTATGCGCTGGGATGCGGACCGGACCATC 72
Db 433 AAGGCTCAGGTTAATCCAGCTATATCCGCTATCTCTGGCATGTCAGAGGCGCAGATT 492
QY 73 CCAGACAGTGCATCTCTGCTCCAGCTCCTGGTGCAGATTCACCTGCCCGCCACACGC 132
Db 493 CCAGATGAGGACATCACAGCTTCCAGTCAGTGCAGTGCAGAGCTCCCAATATGGA 552
QY 133 AGGTGGAGAGCAGTGCAGGGGATGGGCTGTGGTGGCCGAGGGTGTGTTCCCAAG 192
Db 553 AGGCTGGAAGTGCAGAGAGGGGATGAGGCTGTGGTGGCCCTGAGATTCAGATGGAACCTGAT 612

```

```

QY 193 GA---GGAGGAGTACTTGCAGGTGATCTACAACGACTCCACCTGGTGGCTGTGGTGGC 249
Db 613 GACCTGAAGGAGTTTCTGCAGATTGACATGGCACCCCTCCATTTTATCATCTGTGGGG 672
QY 250 ACCAGGGAGCGATCCCGGGGCGCTGGGCAAGAGTCTTCCCGGAGCTACCGGTGGGT 309
Db 673 ACCAGGGGCGCATGCGAGAGGTGATGCAATCGAGTTTGGCCCCCATGTACAAGATCAAT 732
QY 310 TACTCCCGGATGGTGGCGCTGGATGGGCTGGAAGGACCCGCTGGGGTTCAGAGGTGATC 369
Db 733 TACAGTCGGGATGGCACTCGCTGGATCTTGGCGGAACCGTCATGGGAACAGGTGCTG 792
QY 370 TCAGGCAATGAGGACCCCTGAGGAGTGGTGTGAAGGACCTTGGGCCCCCCTGTTGCC 429
Db 793 GATGGAATTAAGAACCCCTATGACATTTTCCCTAAAGGACTTGGAGCGCCCATTTGAGCC 852
QY 430 CGACTGGTTCGCTTTCACCCCGGGCTGACCGGGTCACTGAGCGTCTGTCTGGGGTAGAG 489
Db 853 AGATTTGTCGGTTCATTCAGTCCAGCCAGCACTCCATGAATGTGTGATGAGAGTGAG 912
QY 490 CTCTATGGCTGCTCTGGAGGATGAGTCTGTCTTACCCGCCCTGTGGGGCAGACA 549
Db 913 CTTTACGGCTGTCTGTGGCTAGATGGCTTGGTGTCTTACAATGCTCCAGCTGGCAGCAG 972
QY 550 ATGATTTATCTGAGC-----CGTGTACCTCAAGGACTCCACCTATCACGGACATACC 603
Db 973 TTTGTACTCCCTGGAGGTTTCCATTTATCTGAATGATTTCTCTATGATGAGGTGTT 1032
QY 604 GTGGCGGACTGCAGTATGGGGTCTGGGCGACGTGGCAGATGGTGTGGTGGGTGGAT 663
Db 1033 GGATACAGCATGACAGAAGGG---CTAGGCCAATTCACCGATGGTGTCTGGCTGGAC 1089
QY 664 GACTTTAGGAAGAGTCAAGGAGTGGCGGTCTGGCCAGGCTATGACTATGTGGAGGAGC 723
Db 1090 GATTTACCCAGACCCCATGAATACCACGTGTGGCCCGGCTATGACTATGTGGGCTGGCG 1149
QY 724 AACCAAGTCTTCCAGTGGCTATGAGATGGAGTGGTGTGAGTTTACCGGCTGAGGGCC 783
Db 1150 AAGCAGAGTGGCCCAATGGCTACATTTGAGATCATGTTTGAATTTACCGCATCAGGAAT 1209
QY 784 TTCAGGCTATGCAAGTCCACTGTAACACATGCACACGCTGGGAGCGCTCTGCTGGC 843
Db 1210 TTCACTACCATGAGGTCCACTGCACACACATGTTTGTCTAAAGTGTGAGATCTTTAAG 1269
QY 844 GGGTGGAAATGTCGCTTCCGGCTGGCCCTGCCATGCCCTGGGAGGGGAGCCCATCGCG 903
Db 1270 GAGGTACAGTGTACTTCCGCTCTGAA--GCCAGTGAAGTGGGAACCTAATGCCATTTC 1326
QY 904 CACNACCTAGGGGCAACCTGGGGGACCCAGAGCCCGGCTGTCTCAGTGGCCCTTGGC 963
Db 1327 TTCCTCCCTTCTCTGGATGACGTCAACCCCGAGTGTGCTGCTGAGTGGCTCTCCAC 1386
QY 964 GGCGGTGTGGCTGCTTTCGAGTGCCTTCTCTTTCGCGGGGCGCTGTGTTACTCTTC 1023
Db 1387 CACCAATGGCCAGTGCATCAAGTGTCAATACCATTTTGCAGATACCTGGATGATGTT 1446
QY 1024 AGCGAAATCTCTTCAATCTCTGATGTGGTGAACAATTCCTCTCCGGCACTGGGAGCAC 1083
Db 1447 AGTGAGATCACTTCCCAATCAGATGCTGCAATGTACAACAACCTCTGAAGCCCTGCCACC 1506
QY 1084 TTCCTCCGCGAGCCGCC 1098
Db 1507 TCTCTATGGCACCC 1521

```

## RESULT 6

Q92523  
ID Q92523 standard; cDNA to mRNA; 3157 BP.

AC Q92523;

DT 26-NOV-1995 (first entry)

DE Human colonic adenocarcinoma kinase 2 (CKK-2) cDNA.

KW Mammary carcinoma kinase 10; MCK-10; transmembrane receptor; CKK-2;

KW receptor tyrosine kinase; colonic adenocarcinoma kinase 2; ss.





CC aberrant expression of CCK-2 genes. Engineered cell lines, containing  
 CC recombinant vectors with the present sequence, are useful for producing  
 CC infectious retroviral particles. The cell lines may also be used to  
 CC evaluate and screen drugs involved in CCK-2 activation and regulation.  
 SQ Sequence 3157 BP; 790 A; 802 C; 759 G; 806 T; 759 G; 806 T;

Query Match 27.4%; Score 327.4; DB 1; Length 3157;  
 Best Local Similarity 58.8%; Pred. No. 6.7e-69;  
 Matches 644; Conservative 0; Mismatches 436; Indels 15; Gaps 4;

QY 13 AAGGGACATTTTGTATCTCGCAAGTCGCGCTATGCCCTGGGCATGACGACCGGACCATC 72  
 DB 433 AAGCTCAGGTTATCCAGCTATATCCGCTATCTCTCTGCGCATGTCAGAGGCCAGATT 492  
 QY 73 CCAGACAGTGACATCTCTCTCAGCTCCTGCTGATGATTCACACTGCGCGCCGCCACAGC 132  
 DB 493 CCAGATGAGGACATCACAGCTTCAGTCAGTGTGTCAGAGTCCACAGCTGCCAATATGA 552  
 QY 133 AGTTGGAGACAGTACGAGGGATGGGCTGTGTCGCCCGCAGGGTGGTGTTCCTCCAAAG 192  
 DB 553 AGCTGAGCTCAGAGAAGGGGATGGAGCTGTGTCGCCCTGAGATTCCAGTGGAACTGAT 612  
 QY 193 GA---GGAGGAGTACTTGCAGGTGATCTACACGACTGCACCTGCTGCTGTGGGC 249  
 DB 613 GACCTGAAGAGTTTCTGACAGATTGACTGACACACCTCCATTTATCACTCTGTGGGG 672  
 QY 250 ACCAGGAGCGGATCGCGGGGCTTGGGCAAGGATCTCTCCGGAGCTACCGGCTGCGT 309  
 DB 673 ACCAGGGGCGCATCAGGAGGTGATGCGATGCTGCTGGCCCATGTACAGATCAAT 732  
 QY 310 TACTCCGGGATGGTCCGCGTGGTGGTGGGAGGACCGTGGGGTCCAGGAGTGATC 369  
 DB 733 TACAGTCGGATGGCACTCGCTGGATCTCTTGGCGAACCCGTCATGGGAACAGGTGCTG 792  
 QY 370 TCAGGCAATGAGGACCTGAGGAGTGGTCTGGAAGGACTTGGGCGCCCATGTTGCC 429  
 DB 793 GATGGAAATAGTAACCCCTATGACATTTCTTAAAGGACTTGGAGCGGCCCATGTGAGCC 852  
 QY 430 CGACTGGTGGCTTACCCCGGGGTGACCGGGTCTGAGGCTCTGCTGGGGGTAGAG 489  
 DB 853 AGATTTGTCGGTTCATCTCCAGTCACCGACCACTCCATGAATGTGTATGAGATGGAG 912  
 QY 490 CTCTAGGCTGCTCTGGAGGATGACCTCTGCTTACCGCCCGCTGGGGGAGACACA 549  
 DB 913 CTTTACGGCTGTCTGGCTAGATGCTTGGTGTCTTACAAATGCTCCAGCTGGGAGCAG 972  
 QY 550 ATGATTTATCTGAGGC-----CGTGTACCTCAACGACTCCACCTATGACGAGACATACC 603  
 DB 973 TTTGTACTCCCTGGAGGTTCCATCATTTATCTGAATGATTTCTGTATGATGAGCTGTT 1032  
 QY 604 GTGGCGGACTGACGATGAGGCTGTGGGCCAGCTGGCCAGATGGTGTGGTGGGGCTGGAT 663  
 DB 1033 GGATACAGCATGACAAAGGG---CTAGGCCAATTGACCGATGGTGTGTCTGCGCTGGAC 1089  
 QY 664 GACTTTAGGAAGAGTCAGAGCTGCGGGTCTGGCAGGCTATGACTATGTGGATGGAGC 723  
 DB 1090 GATTTACCCAGACCATGAATACACAGTGTGGCGCGGCTATGACTATGTGGGCTGGGG 1149  
 QY 724 AACACAGCTTCTCCAGTGGCTATGTGGAGATGGAGTTTGTAGTTGACCGGCTGAGGGCC 783  
 DB 1150 AAGGAGAGTGCCACCAATGGCTACATTTAGATCATGTTTGAATTTGACCGCATAGGAAT 1209  
 QY 784 TTCCAGGCTATGAGGTCACCTGTATACAACTGTACACAGCTGGGAGCCGCTGTGCTGCGC 843  
 DB 1210 TTCACTACCATGAAGTCCCATGCAACAACATGTTTGTGTAAGGTGTGAAGATCTTTAAG 1269  
 QY 844 GGGGTGAATGTCGCTTCGGGCTGCGCTGCGCTGCGCTGGAGGGGAGGAGCCCATGGCC 903  
 DB 1270 GAGGTACAGTGTACTTCTCGCTCTGAA---GCCAGTGTAGTGGGAACCTAATGCCATTTCC 1326  
 QY 904 CACAACCTAGGGGGCAACCTGGGGAGCCCGAGAGCCGCGGCTGTCTCAGTGCCTGTCGGC 963

DB 1327 TTCCCCCTTCTCTGGATGACGTCAACCCAGTGTCTCGTTTGTACGGTSCCTCTCCAC 1386  
 QY 964 GGCGTGTGGCTCGCTTCTCTCAGTGGCGCTTCTCTTTTGGGGGCGCTGTGTTACTTTC 1023  
 DB 1387 CACCAATGGGCGAGTGGCCATCAAGTGTCAATACCATTTTGCAGATACCTGGATGATTC 1446  
 QY 1024 AGCGAAATCTCTTCTCATCTCTGATGTGTGAACAATTCCTCTCCGGCACTGGGAGGACC 1083  
 DB 1447 AGTGAGATCACTTCCCAATCAGATGCTGCAATGTACAACTCTGAAGCCCTGCCACC 1506  
 QY 1084 TTCCCGCGACCCGCC 1098  
 DB 1507 TCTCTATGGCACCC 1521

## RESULT 8

V48292  
 ID V48292 standard; cDNA; 3096 BP.  
 AC V48292;  
 DT 16-NOV-1998 (first entry)  
 DE Discoidin domain receptor 2 gene.  
 KW Discoidin domain receptor; transmembrane; metatasis; collagen; ss;  
 KW Cleidocranial dysplasia; Sickler syndrome; extracellular matrix; MMP-1.  
 OS Homo sapiens.  
 FH Key  
 FT CDS Location/Qualifiers  
 FT 354..2921  
 FT /\*tag= a  
 FT /product= "Discoidin domain receptor"  
 FT 354..416  
 FT /\*tag= b  
 FT mat\_peptide 417..2918  
 FT /\*tag= c  
 WO9834954-A2.  
 PD 13-AUG-1998.  
 PF 05-FEB-1998; CA0093.  
 PR 06-FEB-1997; US-041578.  
 PA (MOUN) MOUNT SINAI HOSPITAL CORP.  
 PI Pawson A, Vogel W;  
 DR WPI; 98-447168/38.  
 DR P-PSDB; W77114.  
 DR Novel ligands of discoidin domain receptor tyrosine kinase,  
 PT especially collagen - useful for treating e.g. metastasis,  
 PT cleidocranial dysplasia or Sickler syndrome  
 PS Disclosure; Fig 22a; 115pp; English.  
 CC The DR can be used to identify and evaluate substances which affect DDR  
 CC receptor tyrosine kinase signalling pathways in the cell. Compounds  
 CC which modulate such signalling pathways can be used to alter  
 CC transformation or metastasis in mammals, to treat conditions involving  
 CC structural or functional deregulation of collagens, e.g. Cleidocranial  
 CC dysplasia or Sickler syndrome, conditions requiring modulation of  
 CC extracellular matrix synthesis, degradation or remodelling, or to treat  
 CC conditions needing modulation of MMP-1 expression such as wound healing.  
 SQ Sequence 3096 BP; 762 A; 791 C; 752 G; 791 T;

Query Match 27.4%; Score 327.4; DB 1; Length 3096;  
 Best Local Similarity 58.8%; Pred. No. 6.7e-69;  
 Matches 644; Conservative 0; Mismatches 436; Indels 15; Gaps 4;

QY 13 AAGGGACATTTTGTATCTCGCAAGTGGCGCTATGCCCTGGGCATGACGACCGGACCATC 72  
 DB 417 AAGCTCAGGTTAATCCAGCTATATGCCGCTATCTCTTGGCATGTCCAGGAGCCAGATT 476  
 QY 73 CCAGACAGTGACATCTCTGCTTCCAGCTCTCTGTCAGATTCACACTGCGCGCCGCCACAGC 132  
 DB 477 CCAGATGAGGACATCAGAGCTTCCAGTCACTGTCAGATTCACAGCTCCAAATATGA 536  
 QY 133 AGTTGGAGACAGTACGAGGGGATGGGCTGTGTCGCCCGCAGGCTGGTGTTCCTCCAAAG 192  
 DB 537 AGGCTGAGCTCAGAGAAGGGGATGGAGCTGTGCTGCTGAGATTCCAGTGGAACTGAT 596  
 QY 193 GA---GGAGGAGTACTTGCAGGTGGATCTACAACTGACACTGCTGTGCTGTGGGCGC 249

```

Db 597 GACCTGAAGAGTTTCTCAGATTGACTTGACACCCCTCCATTATATCACTCTGGTGGG 656
QY 250 ACCCAGGACGCGATCGCGGGCTGGCAAGAGTCTCCCGAGCTACCGGCTCGGT 309
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 657 ACCCAGGCGCCATCGAGGAGTTCATGCGATCGATGTTGCCCACTACAAATCAAT 716
QY 310 TACTCCCGGATGTCGCGCTGGATGGGCTGGAGGAGCCCTGGGTGAGAGGTGATC 369
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 717 TACACTCGGATGCGACTCGCTGGATCTCTTGGCGGAACCGCTCATGGAACAGGTGCTG 776
QY 370 TCAGCATAGGACCCCTGAGGAGTGTGCTGAAGGACCTTGGGCCCCCATGTTGCC 429
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 777 GATGAAATAGTAACCCCTATGACATTTTCTAAAGACTTGGAGCGCCCATGTAGCC 836
QY 430 CGACTGGTTCGCTTCTACCCCGGCTGACCGGCTCATGAGCGTGTCTGCGGGTAGAG 489
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 837 AGATTGTCGGTTCATTCAGTACACACACCTCCATGATGTGTGTAGAGTGGAG 896
QY 490 CTCTATGCTGCTCTGAGGAGTACCTCTGCTTTACACCGCCCTGTGGGCGAGACA 549
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 897 CTTTACGGCTGTGCTGCTAGAGTGTGCTTACAAATGCTCCAGCTGGCGCAGCAG 956
QY 550 ATGTATTATCTGAGC-----CGTGTACCTCAACGACTCCACCTATGACGGACATACC 603
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 957 TTTGACTCCCTGGAGGTTCCATCATTTATCTGAATGATTTCTGTATGATGGAGCTGTT 1016
QY 604 GTGGCGGACTCAGTATGCGGCTCTGGCCAGCTGGCAGATGTTGTGGTGGGCTGGAT 663
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1017 GATACAGCATCAGAGAGG---CTAGGCCAATTGACCGATGTGTGTCTGGCTGGAC 1073
QY 664 GACTTTAGGAAGTCAAGAGCTCGGGCTCTGGCAGCTATGACTATGTGGGATGAGC 723
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1074 GATTTACCCAGCACCATGAATACACGTTGGCCCGCTATGACTATGTGGCTGGCGG 1133
QY 724 ACCCAGCTTCTCAGTGGCTATGTGGAGTGGATTTGAGTTGACCGGCTGAGGGCC 783
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1134 AACGAGAGTGCCCAATGGCTACATTTGAGATCATGTTTGAATTTGACCGCATCAGGAAT 1193
QY 784 TTCCAGGCTATCAGGTCCACTGTAAACATGACACGCTGGGAGCCGCTCTGCTGGC 843
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1194 TTCACCTACATGAGTCCACTGCAACACATGTTTGTAAAGTGTGAAGTCTTTAAG 1253
QY 844 GGGGTGAATGTCGTTCCGGCTGGCCCTGCGCATGCGCTGGGAGGGAGCCCATGGCC 903
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1254 GAGGTACAGTGTACTTCCGCTCTGAA---GCCAGTGAAGTGGGAACCTTAATGCCATTTCC 1310
QY 904 CACAACCTAGGGGCAACCTGGGGAGCCCGAGCCCGGCTGTCTAGTCCGCTTGGC 963
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1311 TTCCCGCTTGTCTGGATGACGTCAACCCAGTGTCTGGTTTGTCCAGGTGCGCTCCAC 1370
QY 964 GCGCGTGTGGCTGCTTCTGAGTGGCCGCTTCTCTTTTGGGGGCTGTTACTCTTC 1023
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1371 CACCGAATGCGCAGTCCCATCAAGTGTCAATACCATTTTGCAGATACCTGGATGATTC 1430
QY 1024 AGCGAATCTCTTCTATCTCTGATGTGTGAACAATTCCTCTCCGGCACTGGGAGGAC 1083
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1431 AGTGAGATCACTTCCCAATCAGATGTGCAATGTACAACTCTGAAGCCCTGCCCCACC 1490
QY 1084 TTCCCGCCAGCCCC 1098
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1491 TCTCCTATGGCACCC 1505
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

RESULT 9
V55895
ID V55895 standard; DNA; 3120 BP.
AC V55895;
DT 19-NOV-1998 (first entry)
DE Receptor protein tyrosine kinase (PTK) subtype tyro-10 encoding DNA.
KW PTK; receptor; protein tyrosine kinase; brain tissue; ss.
OS Rattus sp.
FH Key Location/Qualifiers
FT CDS 485..3049

```

```

/*tag= a
/product= *PTK subtype tyro-10*

US5811516-A.
22-SEP-1998.
02-JUN-1995; 456647.
15-MAY-1992; US-884486.
02-MAY-1994; US-237401.
02-JUN-1995; US-456647.
PI (SALK ) SALK INST BIOLOGICAL STUDIES.
Lai CHC, Lemke GE.
WPI: 98-530939/45.
P-PSDB; W79152.
PT Receptor protein tyrosine kinase polypeptide, tyro-3 - preferably
    expressed in brain tissue
PS Example 2; Columns 53-60; 46pp; English.
CC This DNA encodes a novel receptor protein tyrosine kinase (PTK)
    polypeptide subtype tyro-10. The invention provides polynucleotide
    sequences encoding novel PTK polypeptide subtypes tyro-1 to tyro-13.
    The PTK subtypes are found expressed predominantly in the brain tissue.
SQ Sequence 3120 BP; 760 A; 786 C; 830 G; 744 T;

```

```

Query Match 26.8%; Score 321; DB 1; Length 3120;
Best Local Similarity 58.4%; Pred. No. 2.2e-67;
Matches 640; Conservative 0; Mismatches 440; Indels 15; Gaps 4;

QY 13 AAGGGACATTTTGTATCCTGCCAAGTCCGCTATGCCCTGGGCATGACGACCGGACCATC 72
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 548 AAAGCTCAGTTAATCCAGGCATATGCCGCTATCCTCTGGCATGTCAGGAGGCCACATT 607
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 73 CCAGCAGTGCATCTCTCTCCAGCTCCTCTGTCAGATTTCCACTGCCCGCCGACACAGC 132
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 608 CCAGATGAGGACATCATCAGCCTCAAGTCAGTGTGTGAGATTCACGCTCCCAATATGGG 667
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 133 AGCTTGAGGAGCAGTACGCGGATGGGCTGTGTGCCCGCAGGCTGCTTTTCCCAAG 192
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 668 AGCTTGAGTCTCTGAAGAAGAGATGAGGCTGTGTCTCTGAGATTCAGATGCAACCCGAT 727
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 193 GA---GGAGGAGTACTTGCAGGTGATCTACAACGACTGCACCTGGTGGCTCTGTGGGC 249
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 728 GACCTGAAGGAATTTCTGCAGATTGACTTGGCAACCTACACTTTATCACTCTTGTGGGG 787
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 250 ACCCAGGAGCGGATCGCGGGGCTTGGGAAGGAGTCTCTCCGGAGCTACCGGCTGCT 309
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 788 ACCCAGGCGGCTCATCGGGGCTATGTCATTTGCAACCCATGTAAGATCAAC 847
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 310 TACTCCGGATGCTCGCGCTGGATGGCTGGAAGACCGCTGGGGTCAGGAGGTGATC 369
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 848 TACAGTCGGATGGCAGTCTGCTGATCTCTGCGCTAACCGGATGGGAGCAGGTGCTT 907
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 370 TCAGGCAATGAGGACCTCAGGAGTGGTGTCTGAAGGACCTTGGGCCCCCATGTTGCC 429
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 908 GATGGAACAGTAACCTTATGATGATTTCTTGAAGGACTTGGAGCCACCCATCGTCGC 967
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 430 CGACTGTTGCTCTTACCCCCGGCTGACCGGCTCATGAGGCTCTGTCTGGGGTAGAG 489
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 968 AGATTTTTCGCTTATCCCACTGCTGACCTCCATGAACTGTCATGAGGTTGAG 1027
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 490 CTCTATGCTGCTCTGGAGGATGGACTCTCTTCTTACACCGCCCTGTGGGGCAGACA 549
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1028 CTTTATGTTGTCTGGCTAGATGGCTGTGATCTCTACATGCTCCAGCTGGACAGCAG 1087
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 550 ATGTATTATCTGAGC-----CGTGTACCTCAACGACTCCACCTATGACGACATACC 603
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1088 TTTGACTCCCTGGAGCTCCATCATTTATCTGAATGATTTCTGTATGATGAGCTGTT 1147
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 604 GTGGCGGACTGCAAGTATGGGGCTTGGGCCAGCTGGCAGATGGTGTGGTGGGCTGGAT 663
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1148 GGTGTACAGCATGACTGAAGGG---CTAGGCCAGTGTGACTGATGAGTATCCCGGCTGGAT 1204
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 664 GACTTTAGGAAGAGTCAAGGAGTCTGGGCTGTGGCCAGGCTATGACTATGTGGGATGGAGC 723
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1205 GATTTACCCAGACCCATGAATACCAGCTGTGGCTGTGCTATGACTACGTGGGATGGCGG 1264
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

QY 724 AACACAGCTTCTCCAGTGGCTATGTTGGAGATGGAGTTTGAAGTTTACCGGCTGAGGGCC 783  
 DB 1265 AACGAAAGTCTACCAACGGTTTCTTGGATCATGTTTGAATTTGACCGATCAGGAAT 1324  
 QY 784 TTCCAGGCTATGAGTCTCAGTGTACACATGACACAGCTGGAGCGGCTGCTGGC 843  
 DB 1325 TTCTACTACCATGAAGTGCACATGCAACAAATGTTTGTCTAAAGGTGTGAAGATTTTAA 1384  
 QY 844 GGGGTGAATGTCGCTTCCGGCTGGCCCTGCGATGGCTGGAGGGAGGAGCCATGCGC 903  
 DB 1385 GAGTCCAGTGTACTTTCGCTCGGA--AGCCAGGAGTGGGAACCCACTGCTCTCTAC 1441  
 QY 904 CACAACCTTAGGGGCAACCTGGGGACCCAGAGCCCGGGCTCTCTCAGTGGCCCTTGGC 963  
 DB 1442 TTTCCTCCTGGTCTGAGCATGTGAACCCAGTGGCCGGTGTTCACGGTGGCCCTCCAC 1501  
 QY 964 GCGCGTGGCTGCTTTCTGAGTGGCCCTTCTCTTTCGGGGCCCTGGTGTACTCTTC 1023  
 DB 1502 CACCGAATGGCAGTGGCCATCAAGTGCCAAATACCAATTTTGGCGACACGTTGATGTT 1561  
 QY 1024 AGCGAATCTCTTCTATCTCTGATGTGGTGAACAATTCCTCTCCGGCACTGGGAGGCACC 1083  
 DB 1562 AGCGAGATCACTTTCACATCAGATGCTGCAATGTATACAACTCTGGAGCCCTTCCCCACC 1621  
 QY 1084 TTCCGCCACCGCCCC 1098  
 DB 1622 TCTCCTATGGCACCC 1636

RESULT 10  
 V65317  
 ID V65317 standard; DNA; 3120 BP.  
 AC V65317;  
 DT 22-JAN-1999 (first entry)  
 DE Receptor protein tyrosine kinase (PTK) subtype tyro-10 encoding DNA.  
 KW PTK; receptor; protein tyrosine kinase; recombinant; grafting;  
 OS diagnosis; tumour; skin transplant; connective tissue; tyro-10; ss.  
 Rattus sp.  
 FH Key  
 FT 485..3049  
 CDS  
 /tag= a  
 /product= "PTK subtype tyro-10"  
 US5837448-A.  
 PD 17-NOV-1998.  
 PF 02-MAY-1994; 237401.  
 PR 15-MAY-1992; US-884486.  
 PR 02-MAY-1994; US-237401.  
 PA (SALK.) SALK INST BIOLOGICAL STUDIES.  
 PI Lai CHC, Lemke GE;  
 DR WPI; 99-023436/02.  
 PT P-PSDB; W81409.  
 PT Nucleic acids encoding protein tyrosine kinase subtypes - for  
 PT identification of new subtypes and treatment of diseases associated  
 PT with the kinase  
 PS Claim 1; Columns 53-58; 47pp; English.  
 CC This DNA encodes a receptor protein tyrosine kinase (PTK) subtype  
 CC tyro-10. The invention provides sequences V65308 to V65313, V65315, and  
 CC V65317 to V65319 that encode proteins having a tyrosine kinase domain and  
 CC a tissue expression pattern of a receptor PTK subtype selected from  
 CC tyro-1, tyro-2, tyro-3, tyro-4, tyro-5, tyro-6, tyro-8, tyro-10, tyro-11,  
 CC and tyro-12, respectively. The polynucleotides are useful for the  
 CC detection of tyrosine kinase domain sequences and detection of tissue  
 CC expression patterns of PTK subtypes. The cDNAs can also be injected into  
 CC oocytes, the protein expressed, and expression products screened for  
 CC using antibodies against tyrosine kinase epitopes. These subtypes  
 CC sequences can be used for the design of oligonucleotides, for use in  
 CC amplification reactions to isolate other subtype sequences. These  
 CC detection protocols are used in the diagnosis of diseases associated with  
 CC (receptor) PTKs. Recombinant vectors expressing the subtypes can be used  
 CC to treat related diseases e.g. tumours, by introduction of the vectors  
 CC into skin transplants, then grafting these into the connective tissue of  
 CC the dermis, thus specifically targeting tumours as the proteins are

CC released from the matrix.  
 SQ Sequence 3120 BP; 760 A; 786 C; 830 G; 744 T;  
 Query Match 26.8%; Score 321; DB 1; Length 3120;  
 Best Local Similarity 58.4%; Pred. No. 2.2e-67;  
 Matches 640; Conservative 0; Mismatches 440; Indels 15; Gaps 4;  
 QY 13 AAGGACATTTTATCTCTCCAGTGGCTATGTTGGAGATGGAGTTTACCGGCTGAGGGCCATC 72  
 DB 548 AAAGTCTCAGGTTAATCCAGCCATATGCCGTATCTCTTGGCATGTCAGGAGGCCACATT 607  
 QY 73 CCAGACAGTGCACATCTCTGCTTCCAGTCTCTCAGTGTGAGATTCACATGCCGCCCCACAC 132  
 DB 608 CCAGATGAGGACATCAGAGCTCAAGTCAAGTGTGAGATTCACAGCTGCTCAAAATATGG 667  
 QY 133 AGTGTGAGAGCAGTACAGGGGATGGGCTGTGTCGCCCGCCAGAGGTCGGTGTTCCTCCAA 192  
 DB 668 AGGTCTGACTCTGAAGAAGAGATGGAGCTGTGTCTTCCAGATTCAGATTCAGATTCAG 727  
 QY 193 GA---GGAGAGTACTTGCAGTGGATCTACACAGTGTGACCTGTGGTCTGTGGTGGGC 249  
 DB 728 GACCTGAAGGAATTTCTGCAGATTTGACTTCCGAACCTACACTTTATCTCTTGTGGG 787  
 QY 250 ACCCAGGAGCGCATCGCGGGCCCTGGGCAAGAGTTCCTCCCGGAGCTACCGCTGCGT 309  
 DB 788 ACCCAGGAGCGCATCGCGGGCTGATGGATTTGAATTTGCACCCATGACAGATCAAC 847  
 QY 310 TACTCCGGGATGTCGCCCTGGATGGGCTGGAAGACCGCTGGGGTCAGGAGGTGATC 369  
 DB 848 TACAGTCGGGATGGCAGTCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 907  
 QY 370 TCAGGCAATGAGGACCTGAGGAGTGGTCTGGAAGACCTGGGGTCAGGAGGTGATC 429  
 DB 908 GATGGAACAGTAAACCTTTATGATGATTTCTTGAAGACTTGGAGCCACCCATCTGTCG 967  
 QY 430 CGACTGTTGCTCTTACCCCGGCTGACCGGTCATGAGGCTCTGTCTGCGGGTAGAG 489  
 DB 968 AGATTGTTGCTCTTACCCAGTCACTGACCTCCATGAACGTGTGATGAGGGTGTAG 1027  
 QY 490 CTCTATGGTCTCTGAGGAGTGGACTCTCTTCTTACACCGCCCTGTGGGGCAGACA 549  
 DB 1028 CTTTATGGTGTGCTCTGGCTAGATGGCTTGGTATCTTCAATGCTCCAGCTGGACAGAG 1087  
 QY 550 ATGATTTATCTGAGC-----CGTGTACCTCAAGACTCCACCTATGACGGACATACC 603  
 DB 1088 TTTTACTCCTCGGAGGCTCCATCATTTATCTGAATGATTTCTCTATGATGAGGTGT 1147  
 QY 604 GTGGCGGACTGCAGTATGGGGTCTGGGCCAGCTGGCAGTGGTGTGGTGGGGTGGAT 663  
 DB 1148 GGTACAGCATGACTGAAGG---CTAGCCAGTGTGACTGTGAGATATCCGGCTGGAT 1204  
 QY 664 GACTTTAGGAAGTCAAGAGCTGGCGGTCTGCCAGGCTATGACTATGTGGATGGAGC 723  
 DB 1205 GATTTTACCAGACCCCATGAATACCACGTGTGGCTGGCTATGACTACGTGGATGGCGG 1264  
 QY 724 AACCCAGCTTCTCCAGTGGCTATGTTGGAGATGGAGTTTGAAGTTTACCGGCTGAGGGCC 783  
 DB 1265 AACGAAGTGTACCAACCGGTTTCAATTTGAGATCATGTTTGAATTTGACCGAATACGAAT 1324  
 QY 784 TTCAGGCTATGAGGCTCCACTGTAAACATGACACAGCTGGAGCCCGCTGCTGCTGCGC 843  
 DB 1325 TTTACTACCATGAGTTCACCTGCAACACATGTTTGTCTAAAGGTGTGAAGATTTTAA 1384  
 QY 844 GGGGTGAATGCTGCTTCCGGCTGGCCCTGCGATGGCTGGAGGGAGGAGCCCATGCGC 903  
 DB 1385 GAGGTCCAGTGTACTTTCGCTCGGA--AGCCAGGAGTGGGAACCCACTGCTCTCTAC 1441  
 QY 904 CACAACCTTAGGGGCAACCTGGGGACCCAGAGCCCGGGCTGTCTCAGTGGCCCTTGGC 963  
 DB 1442 TTTCCTCCTGGTCTGAGCATGTGAACCCAGTGGCCGGTGTTCACGGTGGCCCTCCAC 1501  
 QY 964 GCGCGTGGCTGCTTTCTGAGTGGCCCTTCTCTTTCGGGGCCCTGGTGTACTCTTC 1023

Db 1502 CACCGAATGGCCAGTGGCCATCAAGTGGCAATACATTTTGGCCAGACGTGGATGATGTC 1561  
 Qy 1024 AGCGAATCTCTTCATCTGATGTGGTGAACAATTCCTCTCGGACACTGGGAGGCACC 1083  
 Db 1562 AGCGAGATCACTTTTCCATCAGATGCTGCAATGTATAACAACCTCTGGAGCCCTTCCACC 1621  
 Qy 1084 TTCGCCGCCAGCCCC 1098  
 Db 1622 TCTCCTATGGACCC 1636

RESULT 11  
 Q64158  
 ID Q64158 standard; DNA; 2128 BP.  
 AC Q64158.  
 DE 03-FEB-1995 (first entry)  
 DE Partial coding sequence of tyrosine kinase receptor protein.  
 KW Tyrosine kinase; receptor; proto-oncogene; trk; detection;  
 KW diagnosis; antibody; treatment; tumour; antisense; ss.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT cds 1..1953  
 FT /\*tag= a  
 FT /product= Partial sequence of protein tyrosine  
 FT kinase.  
 FN DE4239817-A.  
 PN 01-JUN-1994.  
 PD 26-NOV-1992; 239817.  
 PR 26-NOV-1992; DE-239817.  
 PA (CHEM-) CHEMOTHERAPEUTISCHES FORSCHUNG.  
 PI Holtrich U, Ruebsamen-waigmann H, Strebhardt K;  
 DR WPI: 94-184380/23.  
 DR P-PSDB: R54089.  
 PT New protein tyrosine kinase and related nucleic acid - vectors,  
 PT transformed cells, etc., useful for diagnosis and treatment of  
 PT tumours  
 PS Claim 4; Page 8; 9pp; German.  
 CC The gene is related to the trk proto-oncogene. Antibodies against  
 CC the encoded polypeptide are useful for diagnosis and for the  
 CC treatment of tumours. The antibodies may also be radiolabelled or  
 CC coupled to a cytotoxin for destruction of cancer cells. Antisense  
 CC nucleic acid can be used to inhibit gene expression.  
 CC Sequence 2128 BP; 532 A; 556 C; 504 G; 536 T;  
 SQ

Query Match 12.0%; Score 143.2; DB 1; Length 2128;  
 Best Local Similarity 57.0%; Pred. No. 2.9e-25;  
 Matches 303; Conservative 0; Mismatches 223; Indels 6; Gaps 2;

Qy 567 CGTGACTCAACGACTCCACCTATGACGACATACCTGGCGGACTGCGAGTATGGGG 626  
 Db 12 CATTTATCTGAATGATTCTGCTATGATGGAGCTTGTGATACAGCATGACAGAAGGG-- 70  
 Qy 627 TCTGGCCAGCTGGCAGATGTTGGTGGGTGGATGACTTTAGGAGAGTCAAGAGCT 686  
 Db 70 -CTAGGCCAATTGACCGATGTTGCTGCGCTGGACGATTTCACCCAGACCCATGAATA 128  
 Qy 687 GCGGTGTGGCGAGCTATGACTATGTGGGATGGAGCAACACAGCTTCTCCAGTGCTA 746  
 Db 129 CCAGTGTGGCCCGCTATGACTATGTGGGTGGCGGAGCAGAGAGTGCACCATGCTA 188  
 Qy 747 TGTGAGATGAGGTTGAGTTTACCGGCTGAGGGCCCTTCCAGGCTATGCAAGTCCACTG 806  
 Db 189 CATTGAGATCATGTTTGAATTTGACCGCATCAGGAATTTCACTACCATGAAGTCCACTG 248  
 Qy 807 TAACACATGCACGCTGGGAGCCGCTGCTGCGCGGGTGGATGCTGCTCCGCGG 866  
 Db 249 CAACAACATGTTTGAAGGTGTGAAGATCTTTAAGGAGGTACAGTCTACTTCCGCTC 308  
 Qy 867 TGGCCCTGCGATGGCTGGGAGGGGAGCCCATGCGCCACACACCTAGGGGGCAACCTGGG 926  
 Db 309 TGAA--GCACGTGAGTGGACCTTAATCCCATTTCCCTTCCCTTGTCTGGATGAGCT 365

Qy 927 GGACCCAGAGCCCGGGCTGTCTCAGTCCCTTGGGGCCGTGTGGCTCTCTTCTGCA 986  
 Db 366 CAACCCAGTGTCTGGTTGTACGGTCCCTCTCCACCACCAATGGCCAGTGCATCAA 425  
 Qy 987 GTGCGCTCTCTTTTGGGGCCCTGTACTTCTAGCGAATCTCTTCACTCTCTGA 1046  
 Db 426 GTGTCAATACCATTTTGCAGATACCTGTGATGTCTAGTGTAGATCACCCTTCCAATCAGA 485  
 Qy 1047 TGTGTGAACATTCCTCTCCGGCACTGGGAGGACCTTCCCGCCAGCCGCC 1098  
 Db 486 TCTGCAATGTACAACTCTGAAGCCCTGCCACCTCTCTTATGGCACCC 537

RESULT 12  
 V23288  
 ID V23288 standard; DNA; 4451 BP.  
 AC V23288;  
 DT 17-AUG-1998 (first entry)  
 DE Synthetic human Factor-VIII gene lacking central B domain.  
 KW Factor-VIII; blood clotting; human; synthetic gene; codon usage; ss.  
 OS Homo sapiens.  
 OS Synthetic.  
 PN WO9812207-A1.  
 PD 26-MAR-1998.  
 PF 18-SEP-1997; U16639.  
 PR 20-SEP-1996; US-717294.  
 PA (GEO) GEN HOSPITAL CORP.  
 PI Haas J, Seed B;  
 DR WPI: 98-217200/19.  
 DR New synthetic eukaryotic gene(s) - in which non-preferred or less  
 PT preferred codon(s) are replaced to provide high level expression in  
 PT mammalian cell(s)  
 PS Claim 20; Fig 13; 92pp; English.  
 CC This synthetic gene codes for a human Factor-VIII protein that  
 CC lacks the central B domain (amino acids 760-1639) of the native  
 CC protein. In the synthetic gene, non-preferred or less preferred  
 CC codons of the native gene (see V2339) are replaced by codons  
 CC favored by highly expressed human genes to provide high-level  
 CC expression in mammalian cells. The synthetic gene was assembled  
 CC from 29 pairs of oligonucleotides (see V23340-97) which served as  
 CC PCR templates. Synthetic genes of the invention (see also  
 CC V23289-91) are used for production of recombinant proteins in  
 CC mammalian cells at levels of at least 500% of those obtained using  
 CC the natural genes. They can also be used in gene therapy. An  
 CC expression vector comprising a synthetic gene and a mammalian cell  
 CC harboring a synthetic gene are also claimed.  
 CC Sequence 4451 BP; 972 A; 1568 C; 1234 G; 677 T;  
 SQ

Query Match 6.2%; Score 73.8; DB 1; Length 4451;  
 Best Local Similarity 50.1%; Pred. No. 9.8e-09;  
 Matches 211; Conservative 0; Mismatches 207; Indels 3; Gaps 1;

Qy 33 CAAGTCCCGCTATGCCCTGGGCATGACGACCGGACCATCCAGACAGTACATCTCTGC 92  
 Db 3959 CAGCTCAGCATGCGCCTGGGCATGGAGAGCAAGGCCATCAGCGACGCCCATCACCGC 4018  
 Qy 93 TTCACGCTCTGTGTGATGATTCACCTGCCCGCCAGCAGAGTGTGGAGAGCATGACGG 152  
 Db 4019 CTCACGCTACTTCAACCAACATGTTGCCACCTGGAGCCCGCCAGCCCTGCACCT 4078  
 Qy 153 GGATGGGGCTGTGTCGCCGAGGGTGGTGTTCCTCCAGGAGGAGGATCTTCAGGT 212  
 Db 4079 GCAGGGCCCGCAGCAGCAGCGCTGGCGCCCGCCAGGTGAACACCCCAAGGCTGCTCAGST 4138  
 Qy 213 GGATCTACACGACTGCACCTGTGTGCTGTGGGACCCAGGAGCGATCTCCGGGGG 272  
 Db 4139 GGAATCCAGAAACCATGAAGTGACTGGGTGACCCAGCCAGGGCGTCAAGAGCTG-- 4197  
 Qy 273 CTTGGGCAAGGAGTTCTCCCGGAGCTTACCGGTGTGCTTACTCCCGGAGTGTGCGCGCTG 332  
 Db 4197 -CTGACGAGCATGTACGTGAAGGAGTTCTGTATCAGCAGCAGCAGGACCGCCACCATG 4255





GenCore version 4.5  
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 4, 1999, 07:31:19 ; Search time 1266.11 Seconds  
(without alignments)  
3006.714 Million cell updates/s

Title: US-08-170-558-7  
 Perfect score: 1197  
 Sequence: 1 GATGCTGACATGAAGGGACA.....CCGAGGGGAGCCCGACCGCC 1197

Scoring table: IDENTITY\_NUC  
Searched: 679419 seqs, 1590154680 residues

```

Database :
GenEmbl.*
1: gb_ba1.*
2: gb_ba2.*
3: gb_om.*
4: gb_ov.*
5: gb_pat.*
6: gb_ph.*
7: gb_pl1.*
8: gb_pl2.*
9: gb_pr1.*
10: gb_pr2.*
11: gb_pr3.*
12: gb_ro.*
13: gb_st.*
14: gb_sts.*
15: gb_sy.*
16: gb_un.*
17: gb_vi.*
18: em_fun.*
19: em_htg.*
20: em_hum1.*
21: em_hum2.*
22: em_in.*
23: em_om.*
24: em_or.*
25: em_ov.*
26: em_pat.*
27: em_ph.*
28: em_pl.*
29: em_ro.*
30: em_sts.*
31: em_sy.*
32: em_un.*
33: em_vi.*
34: gb_htg1.*
35: gb_htg2.*
36: gb_in1.*
37: gb_in2.*
38: em_ba1.*
39: em_ba2.*
40: em_hum3.*
41: em_hum4.*
42: gb_pr4.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1197	100.0	3637	5 I80845	I80845 Sequence 3

[illegible]

## ALIGNMENTS

```

RESULT      1
180845
180845 LOCUS
180845 DEFINITION
180845 Sequence 3 from patent US 5709858.
180845 ACCESSION
93209135
NID
180845.1 VERSION
GI:3209135
KEYWORDS
SOURCE
Unknown.
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 3637)
AUTHORS
Godowski, P.J., Mark, M.R. and Scadden, D.T.
TITLE
Antibodies specific for Rse receptor protein tyrosine kinase
JOURNAL
Patent: US 5709858-A 3 20-JAN-1996;
FEATURES
Location/Qualifiers
1..3637
/organism="unknown"
BASE COUNT
721 a 1094 c 1079 g 743 t

```

BASE COUNT	721 a	1094 c	1079 g	743 t
ORIGIN				

```

Query Match          Score 1197;  DB 5;  Length 3637;
Best Local Similarity 100.0%;  Pred. No. 6.7e-249;

```

Matches 1197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATCCTGACATGAAGGACATTTTATCTTCCCAAGTGGCGCTATGCTTGGCGATGCGAG 60  
Db 136 GATCCTGACATGAAGGACATTTTATCTTCCCAAGTGGCGCTATGCTTGGCGATGCGAG 195  
Qy 61 GACCGGACCATCCAGACAGTACATCTCTGCTTCCAGTCTCTGCTGAGATTCACATGCC 120  
Db 196 GACCGGACCATCCAGACAGTACATCTCTGCTTCCAGTCTCTGCTGAGATTCACATGCC 255  
Qy 121 GCCCGCACAGCAGTGGAGAGCAGTACGGGATGGGGCTGGTGGCCCGCAGGGTGG 180  
Db 256 GCCCGCACAGCAGTGGAGAGCAGTACGGGATGGGGCTGGTGGCCCGCAGGGTGG 315  
Qy 181 GTGTTTCCCAAGGAGGAGTACTTCCAGTGGATCTACAAAGTCTGACCTGGTGGCT 240  
Db 316 GTGTTTCCCAAGGAGGAGTACTTCCAGTGGATCTACAAAGTCTGACCTGGTGGCT 375  
Qy 241 CTGTTGGGACCCAGGACGSCATGCCGGGGCTGGGCAAGGAGTCTCCCGGAGCTAC 300  
Db 376 CTGTTGGGACCCAGGACGSCATGCCGGGGCTGGGCAAGGAGTCTCCCGGAGCTAC 435  
Qy 301 CGGTGCTTACTCCCGGATGCTGGCGCTGGATGGGCTGGAGGACCGCTGGGCTAG 360  
Db 436 CGGTGCTTACTCCCGGATGCTGGCGCTGGATGGGCTGGAGGACCGCTGGGCTAG 495  
Qy 361 GAGGTGATCTCAGCAATGAGGACCTGAGGAGTGGTCTGAAGGACCTTGGGCCCCC 420  
Db 496 GAGGTGATCTCAGCAATGAGGACCTGAGGAGTGGTCTGAAGGACCTTGGGCCCCC 555  
Qy 421 ATGTTGCCCGACTGGTTCCTTACCCCGGCTGACCGGCTGATGAGCGTCTGCTG 480  
Db 556 ATGTTGCCCGACTGGTTCCTTACCCCGGCTGACCGGCTGATGAGCGTCTGCTG 615  
Qy 481 CGGTGAGCTCTATGCTGCTCTGAGGATGGATCTCTGCTTACACCGCCCTGTTG 540  
Db 616 CGGTGAGCTCTATGCTGCTCTGAGGATGGATCTCTGCTTACACCGCCCTGTTG 675  
Qy 541 GGGCAGACAATGATTTATCTGAGCGCTGTACCTCAACGACTCACCTATGACGGACAT 600  
Db 676 GGGCAGACAATGATTTATCTGAGCGCTGTACCTCAACGACTCACCTATGACGGACAT 735  
Qy 601 ACCGTGGCGGAGTGCAGTATGGGGTCTGGGCGAGCTGGCAGATGCTGTTGGGGCTG 660  
Db 736 ACCGTGGCGGAGTGCAGTATGGGGTCTGGGCGAGCTGGCAGATGCTGTTGGGGCTG 795  
Qy 661 GATGACTTTAGAGAGTCAGAGCTCGGGCTGCGCCAGCTATGACTATGTTGGGATGG 720  
Db 796 GATGACTTTAGAGAGTCAGAGCTCGGGCTGCGCCAGCTATGACTATGTTGGGATGG 855  
Qy 721 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGGAGTTTGGATTTGACCGGCTGAGG 780  
Db 856 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGGAGTTTGGATTTGACCGGCTGAGG 915  
Qy 781 GCTTTCCAGGATGACAGTCCACTGTAAACATGACACGCTGGAGCCGCTCTGCCT 840  
Db 916 GCTTTCCAGGATGACAGTCCACTGTAAACATGACACGCTGGAGCCGCTCTGCCT 975  
Qy 841 GCGGGGTGGAATGCTGCTTCCGGCTGCGCTGCCATGGCTGGGAGGGAGCCCATG 900  
Db 976 GCGGGGTGGAATGCTGCTTCCGGCTGCGCTGCCATGGCTGGGAGGGAGCCCATG 1035  
Qy 901 GCGCAACACCTAGGGGCAACCTGGGGGACCCAGAGCCCGGCTGTCTCAGTGGCCCTT 960  
Db 1036 GCGCAACACCTAGGGGCAACCTGGGGGACCCAGAGCCCGGCTGTCTCAGTGGCCCTT 1095  
Qy 961 GCGGCGGTGGCTGCTTCTGAGTGGCGCTTCTCTTCTTGGGGCCCTGTTACTC 1020  
Db 1096 GCGGCGGTGGCTGCTTCTGAGTGGCGCTTCTCTTCTTGGGGCCCTGTTACTC 1155  
Qy 1021 TTCAGGAAATCTCTCTCATCTGATGTTGTAACAAATTCCTCTCGGCACTGGGAGGC 1080  
Db 1156 TTCAGGAAATCTCTCTCATCTGATGTTGTAACAAATTCCTCTCGGCACTGGGAGGC 1215

Qy 1081 ACCTTCCCGCAGCCCCCTGTGTGGCGCTGGCCACACTCCCACTTCCAGCAGCTTG 1140  
Db 1216 ACCTTCCCGCAGCCCCCTGTGTGGCGCTGGCCACACTCCCACTTCCAGCAGCTTG 1275  
Qy 1141 GAGCTGGAGCCCGCAGGCGCAGCAGCCCGTGGCCAAAGCCCGAGGGAGCCCGACGCC 1197  
Db 1276 GAGCTGGAGCCCGCAGGCGCAGCAGCCCGTGGCCAAAGCCCGAGGGAGCCCGACGCC 1332

RESULT 2  
180847 I80847 1197 bp DNA PAT 20-MAR-1998  
LOCUS Sequence 7 from patent US 5709858.  
DEFINITION I80847  
ACCESSION I80847  
NID 93209137  
VERSION I80847.1 GI:3209137  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1197)  
AUTHORS Godowski, P.J., Mark, M.R. and Scadden, D.T.  
TITLE Antibodies specific for Rse receptor protein tyrosine kinase  
JOURNAL Patent: US 5709858-A 7 20-JAN-1998;  
FEATURES Location/Qualifiers  
1. 1197  
source /organism="unknown"  
BASE COUNT 199 a 351 c 407 g 240 t  
ORIGIN

Query Match 100.0%; Score 1197; DB 5; Length 1197;  
Best Local Similarity 100.0%; Pred. No. 7.6e-249;  
Matches 1197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GATGTGACATGAAGGACATTTTATCTTCCCAAGTGGCGCTATGCTTGGCGATGCGAG 60  
Db 1 GATGTGACATGAAGGACATTTTATCTTCCCAAGTGGCGCTATGCTTGGCGATGCGAG 60  
Qy 61 GACCGGACCATCCAGACAGTACATCTCTGCTTCCAGTCTCTGCTGAGATTCACACTGCC 120  
Db 61 GACCGGACCATCCAGACAGTACATCTCTGCTTCCAGTCTCTGCTGAGATTCACACTGCC 120  
Qy 121 GCCCGCACAGCAGTGGAGAGCAGTGAACGGGATGGGGCTGGTGGCCCGCAGGGTGG 180  
Db 121 GCCCGCACAGCAGTGGAGAGCAGTGAACGGGATGGGGCTGGTGGCCCGCAGGGTGG 180  
Qy 181 GTGTTTCCCAAGGAGGAGTACTTGCAGTGGATCTACACGACTGCACCTGGTGGCT 240  
Db 181 GTGTTTCCCAAGGAGGAGTACTTGCAGTGGATCTACACGACTGCACCTGGTGGCT 240  
Qy 241 CTGTTGGGACCCAGGACGCGATCCGGGGCTGGGCAAGGAGTCTCCCGGAGCTAC 300  
Db 241 CTGTTGGGACCCAGGACGCGATCCGGGGCTGGGCAAGGAGTCTCCCGGAGCTAC 300  
Qy 301 CGGCTGCTTACTTCCCGGATGGTCCCGCTGGATGGGCTGGAAAGGACCGCTGGGGTGG 360  
Db 301 CGGCTGCTTACTTCCCGGATGGTCCCGCTGGATGGGCTGGAAAGGACCGCTGGGGTGG 360  
Qy 361 GAGGTGATCTCAGGCAATGAGGACCGCTGAGGAGTGGTGGTGAAGGACCTTGGGGCCCCC 420  
Db 361 GAGGTGATCTCAGGCAATGAGGACCGCTGAGGAGTGGTGGTGAAGGACCTTGGGGCCCCC 420  
Qy 421 ATGTTTGGCGGACTGGTTCGCTTCTACCCCGGCTGACCGGCTCATGAGCGTCTGCTG 480  
Db 421 ATGTTTGGCGGACTGGTTCGCTTCTACCCCGGCTGACCGGCTCATGAGCGTCTGCTG 480  
Qy 481 CGGGAAGAGCTCTATGGCTGCTCTGGAGGATGAGACTCCTCTCTTACACCCCGCTGTG 540  
Db 481 CGGGAAGAGCTCTATGGCTGCTCTGGAGGATGAGACTCCTCTCTTACACCCCGCTGTG 540  
Qy 541 GGGCAGACAATGATTTATCTGAGCGCTGTACCTCAACGACTCCACCTATGACGGACAT 600



Db	541	GGGAGACATGATTTATCTGAGCGCTGTACTCAACGACTCCACCTATGACGACAT	600
Qy	601	ACCGTGGCGGACTGCAGTATGGGGTCTTGGCGAGCTGGCAGATGCTGTGGGGCTG	660
Db	601	ACCGTGGCGGACTGCAGTATGGGGTCTTGGCGAGCTGGCAGATGCTGTGGGGCTG	660
Qy	661	GATGACATTTAGGAAGATCAGGAGCTCGGGTCTGGCGAGCTGATGACTATGAGGATG	720
Db	661	GATGACATTTAGGAAGATCAGGAGCTCGGGTCTGGCGAGCTGATGACTATGAGGATG	720
Qy	721	AGCAACACAGCTTCTCAGTGGCTATGAGATGAGTGTGAGTTGACCGGCTGAGG	780
Db	721	AGCAACACAGCTTCTCAGTGGCTATGAGATGAGTGTGAGTTGACCGGCTGAGG	780
Qy	781	GCCTTCCAGGCTATGAGGCTCACTGTAACAACATGACAGCTGGGAGCCCTGCTGCCT	840
Db	781	GCCTTCCAGGCTATGAGGCTCACTGTAACAACATGACAGCTGGGAGCCCTGCTGCCT	840
Qy	841	GGCGGGTGGAAATCTCGCTTCCGGCTGGCCCTGATGAGCTGGGAGGGAGCCCATG	900
Db	841	GGCGGGTGGAAATCTCGCTTCCGGCTGGCCCTGATGAGCTGGGAGGGAGCCCATG	900
Qy	901	CGCCACACCTAGGAGGCACTGGGGAGCCCGAGAGCCGGCTGCTCAGTGGCCCTT	960
Db	901	CGCCACACCTAGGAGGCACTGGGGAGCCCGAGAGCCGGCTGCTCAGTGGCCCTT	960
Qy	961	GGCGGGCTGTGGCTTCTTCTCAGTGGCGCTTCTTCTTGGGGGGCCCTGGTACTC	1020
Db	961	GGCGGGCTGTGGCTTCTTCTCAGTGGCGCTTCTTCTTGGGGGGCCCTGGTACTC	1020
Qy	1021	TTCAGCGAAATCTCTTCTATCTGTATGTGTGTAACAATTCCTCTCGGCACTGGGAGGC	1080
Db	1021	TTCAGCGAAATCTCTTCTATCTGTATGTGTGTAACAATTCCTCTCGGCACTGGGAGGC	1080
Qy	1081	ACCTTCCCGCAGCCCTTGTGGCGCTTGGCCCACTCCACCACTTCAGCAGCTTG	1140
Db	1081	ACCTTCCCGCAGCCCTTGTGGCGCTTGGCCCACTCCACCACTTCAGCAGCTTG	1140
Qy	1141	GAGCTGGAGCCAGAGGCGAGCCGCTGGCCCAAGCCGAGGGAGCCCGACGCC	1197
Db	1141	GAGCTGGAGCCAGAGGCGAGCCGCTGGCCCAAGCCGAGGGAGCCCGACGCC	1197
RESULT	3		
HUMCAK			
LOCUS	HUMCAK	3803 bp	mRNA
DEFINITION	Homo sapiens tyrosine protein kinase (CAK) gene, complete cds.		
ACCESSION	L20817		
NID	g306474		
VERSION	L20817.1	GI:306474	
KEYWORDS	A5-antigen; factor V; factor VIII; tyrosine protein kinase.		
SOURCE	Homo sapiens fetus liver cDNA to mRNA.		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 3803)		
	Perez, J.L., Shen, X., Finkernagel, S., Sciorra, L., Jenkins, N.A., Gilbert, D.J., Copeland, N.G. and Wong, T.W.		
	Identification and chromosomal mapping of a receptor tyrosine kinase with a putative phospholipid binding sequence in its ectodomain		
JOURNAL	Oncogene 9, 211-219 (1994)		
MEDLINE	94134417		
REFERENCE	2 (bases 1 to 3803)		
AUTHORS	Johnson, J.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (17-FEB-1993)		
REFERENCE	3 (sites)		
AUTHORS	Johnson, J.D., Edman, J.C. and Rutter, W.J.		
TITLE	A receptor tyrosine kinase found in breast carcinoma cells has an extracellular discoidin I-like domain		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 90, 5677-5681 (1993)		

MEDLINE	93296201		
REFERENCE..	4 (sites)		
AUTHORS	Wong, T.		
TITLE	Direct Submission		
JOURNAL	Submitted (06-JUL-1993)		
FEATURES	Location/Qualifiers		
source	1..3803		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/dev_stage="fetus"		
	/tissue_type="liver"		
	/map="chromosome 6"		
	1..2934		
	/gene="CAK"		
gene	1..192		
5'UTR	193..246		
sig_peptide	193..246		
CDS	193..2934		
	/gene="CAK"		
	/note="N-linked glycosylation sites: bps 829-831, 976-978, 1306-1308, 1378-1380; domains: transmembrane-bps. 1441-1521, kinase-bps. 2041-2877; furin cleavage site: bps. 1102-1113"		
	/codon_start=1		
	/product="tyrosine protein kinase"		
	/protein_id="AA18019.1"		
	/db_xref="PID:g306475"		
	/db_xref="GI:306475"		
	/translation="MGPEALSSLLLLLVASGDADMKGFDPKACRYALGMDRTIPDSDIASSSWSSTAAHRSLESSDGDGAWCPAGSVFPKEEYIQVDLQRLHLVALVGTQRRHAGGLGKFSRSLRYSRDGRWGMKDRWQEVISGNEDEPGVLDLGPVAVARLVFPRADRVMSVCLRVLYGLWRDGLLSYAPVQTMVLSAEVYLDSTYDGH TVGGLGYGLQGLADVVDLDRKSOELRWVPGYDVVGNHSSFSGYSVEMEFDFR LRAQAMQVHCNMHTLGARLPGCGECFRFRGPMAMWEGPMHNLGNLGDPRARV SVPLGVRVARELQCFRLFAGFWLLFSEISFISDVVNSSPALGGTFFPAWPPPPPP TSFSSLELEPRGQQPVAKASPTAILGLVAIILLILLIALLMLRLHRLVLSKAE RRYLEELTVHLSVPGDFTILINRRPGPREPPYQEPKPNPSPHSPACVPPNLSALL SNPAYRLLATYARPPRGPGPPTPAWAKPTNTQAYSGDYMEPEKPGAPLLPQPNSV PHYAEADIVTLOGVTGNTVAVPALPGAVDGPVDFPRSLRFRFKELGQGFQEV HLCVDSQDVLSDFLNVRKHPHLLVAVKILRPDATKNARNDFLAEVKIMSLKDP NIIRLLGVCVDDPLCMITDMENGLNQFLSAHQEDKRAAGAPGQGAQGTIISY PMLLHVAAQASGMWECILMGKFTASDVWAFVTLMEVLMCLRAQPFQGLTDEQVIER VQGRAVLPRIRMAWECILMGKFTASDVWAFVTLMEVLMCLRAQPFQGLTDEQVIER AGEFFRDQGRVLYLSRPPACPGQGLYELMLRCWSRESRQRPFFSQLHFLAEDALNTV"		
BASE COUNT	721 a 1184 c 1123 g 775 t		
ORIGIN			
Query Match	99.9%	Score 1195.4;	DB 9; Length 3803;
Best Local Similarity	99.9%;	Pred. No. 1.5e-248;	
Matches 1196;	Conservative 0;	Mismatches 1;	Indels 0; Gaps 0;
Qy	1	GATGCTGACATGAAGGACATTTTATCTGCCAAGTCCGCTATGCTCCCTGGGCATGCAG	60
Db	247	GATGCTGACATGAAGGACATTTTATCTGCCAAGTCCGCTATGCTCCCTGGGCATGCAG	306
Qy	61	GACCGACCATCCAGACAGTGCACATCTCTCTCCAGCTCTGCTGCAGATTCCTCCACTGCC	120
Db	307	GACCGACCATCCAGACAGTGCACATCTCTCTCCAGCTCTGCTGCAGATTCCTCCACTGCC	366
Qy	121	GCCCGCCACAGCAGGTTGGAGACAGTGCACGGGATGGGGCTGGTCCCGCCGAGGGTCG	180
Db	367	GCCCGCCACAGCAGGTTGGAGACAGTGCACGGGATGGGGCTGGTCCCGCCGAGGGTCG	426
Qy	181	GTGTTTCCCAAGAGAGGAGTACTTTGAGGTGATCTACACAGTGCACCTGCTGGCT	240
Db	427	GTGTTTCCCAAGAGGAGTACTTTGAGGTGATCTACACAGTGCACCTGCTGGCT	486
Qy	241	CTGGTGGGACCCAGGAGGAGGATGCCGGGGCCCTGGCGAGGATTTCTCCCGGAGCTAC	300
Db	487	CTGGTGGGACCCAGGAGGAGGATGCCGGGGCCCTGGCGAGGATTTCTCCCGGAGCTAC	546

QY 301 CGGCTGCGTTACTCCGGGATGGTCCGCGCTGGATGGGCTGAAGGACCGCTGGGTCAG 360  
Db 547 CGGCTGCGTTACTCCGGGATGGTCCGCGCTGGATGGGCTGAAGGACCGCTGGGTCAG 606  
QY 361 GAGGTGATCTCAGGAATGAGGACCTGAGGAGTGGTCTGAAGGACCTTGGGCCCCC 420  
Db 607 GAGGTGATCTCAGGAATGAGGACCTGAGGAGTGGTCTGAAGGACCTTGGGCCCCC 666  
QY 421 ATGGTTGCCGACTGGTTCGCTTACCCCCGGGCTGACCCGGTCAATGAGGCTGTGCTG 480  
Db 667 ATGGTTGCCGACTGGTTCGCTTACCCCCGGGCTGACCCGGTCAATGAGGCTGTGCTG 726  
QY 481 CGGTAGAGCTATGGCTGCTCTGGAGGATGAGTCCCTGTCTTACACCGCCCTGTG 540  
Db 727 CGGTAGAGCTATGGCTGCTCTGGAGGATGAGTCCCTGTCTTACACCGCCCTGTG 786  
QY 541 GGGCAGACAATGATTTATCTCAGGCGCTGTACCTCAACGACTCCACCTATGACGGACAT 600  
Db 787 GGGCAGACAATGATTTATCTCAGGCGCTGTACCTCAACGACTCCACCTATGACGGACAT 846  
QY 601 ACCGTGGGGGACTCAGATGATGGGGTCTGGGCCAGCTGGCAGATGTTGTTGGGGCTG 660  
Db 847 ACCGTGGGGGACTCAGATGATGGGGTCTGGGCCAGCTGGCAGATGTTGTTGGGGCTG 906  
QY 661 GATGACTTTAGGAAGAGTCAGGAGCTGGGCTCTGCCAGCTATGACTATGTTGGGATG 720  
Db 907 GATGACTTTAGGAAGAGTCAGGAGCTGGGCTCTGCCAGCTATGACTATGTTGGGATG 966  
QY 721 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGTGGATTTGATTTGACCGCTGAGG 780  
Db 967 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGTGGATTTGATTTGACCGCTGAGG 1026  
QY 781 GCCTTCCAGGCTATCAGGTCACCTGTAAACACATGACACGCTGGGAGCCCTGTGCT 840  
Db 1027 GCCTTCCAGGCTATCAGGTCACCTGTAAACACATGACACGCTGGGAGCCCTGTGCT 1086  
QY 841 GCGGGGTGGAATGTCGCTTCCGGGCTGGCCCTGCCATGGCTGGGAGGGAGCCCATG 900  
Db 1087 GCGGGGTGGAATGTCGCTTCCGGGCTGGCCCTGCCATGGCTGGGAGGGAGCCCATG 1146  
QY 901 GCCCAACACCTAGGGGGCAACCTGGGGGACCCCGAGAGCCCGGCTGTCTCAGTGGCCCTT 960  
Db 1147 GCCCAACACCTAGGGGGCAACCTGGGGGACCCCGAGAGCCCGGCTGTCTCAGTGGCCCTT 1206  
QY 961 GCGGCGGTGGCTGCTGCTTCTCAGTGGCGCTTCTCTTTGGGGGCGCTGTGTTACTC 1020  
Db 1207 GCGGCGGTGGCTGCTGCTTCTCAGTGGCGCTTCTCTTTGGGGGCGCTGTGTTACTC 1266  
QY 1021 TTCAGGAAATCTCTTCTCATCTGATGTTGAGCAATTCCTTCCGGCACTGGGAGGC 1080  
Db 1267 TTCAGGAAATCTCTTCTCATCTGATGTTGAGCAATTCCTTCCGGCACTGGGAGGC 1326  
QY 1081 ACCTTCCGCGCCAGCCCTGTTGGCGGCTGGCCACCTCCACCACTTCAGCAGCTTG 1140  
Db 1327 ACCTTCCGCGCCAGCCCTGTTGGCGGCTGGCCACCTTCACCACTTCAGCAGCTTG 1386  
QY 1141 GAGCTGGAGCCAGAGGCGCAGGAGCCCTGGCCAGCCCGAGGGAGCCCGACCGCC 1197  
Db 1387 GAGCTGGAGCCAGAGGCGCAGGAGCCCTGGCCAGCCCGAGGGAGCCCGACCGCC 1443

## RESULT 4

HSRETYK1 HSRETYK1 3841 bp mRNA PRI 20-APR-1995  
LOCUS H.sapiens EDDR1 gene for receptor tyrosine kinase.  
DEFINITION  
ACCESSION Z29093  
NID 9732799  
VERSION Z29093.1 GI:732799  
KEYWORDS receptor tyrosine kinase.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;  
Primates; Catarrhini; Homidae; Homo.

## REFERENCE

AUTHORS Laval,S., Butler,R., Shelling,A.N., Hanby,A.M., Poulsom,R. and Gansan,T.S.  
TITLE Isolation and characterization of an epithelial-specific receptor tyrosine kinase from an ovarian cancer cell line  
JOURNAL Cell Growth Differ. 5 (11), 1173-1183 (1994)  
MEDLINE 95151638  
REFERENCE 2 (bases 1 to 3841)  
AUTHORS Shelling,A.N., Butler,R., Jones,T., Laval,S., Boyle,J.M. and Gansan,T.S.  
TITLE Localization of an epithelial-specific receptor kinase (EDDR1) to chromosome 6q16  
JOURNAL Genomics 25 (2), 584-587 (1995)  
MEDLINE 95309932  
REFERENCE 3 (bases 1 to 3841)  
AUTHORS Keding,C.  
TITLE Direct Submission  
JOURNAL Submitted (17-DEC-1993) Claude Keding, CNRS Laboratoire de genetique moleculaire-U184, INSERM, 11, rue Humann, Strasbourg, Alsace, 67085 cedex, FRANCE  
FEATURES  
source  
1. .3841  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/clone="RTK6"  
/dev\_stage="Adult"  
/tissue\_type="Ovary"  
/cell\_type="Epithelial cell"  
/cell\_line="SKOV-3"  
/clone\_lib="pks"  
336..423  
/note="putative signal peptide"  
337..2967  
/gene="EDDR1"  
337..2967  
/gene="EDDR1"  
/standard\_name="epithelial discoidin domain receptor"  
/product="receptor tyrosine kinase"  
/protein\_id="CAA82335.1"  
/db\_xref="PID:g732800"  
/db\_xref="GI:732800"  
/db\_xref="SWISS-PROT:Q08345"  
/translation="MGPEALSLLLLLLVASGDADMKGHFPAKRYALGMDRTIPD SDIASSSSDSTAARHSRLSSDGDGAWCPAGSVFPEEYLOVDLQRLHLVALVGT QGRHAGGLGKEFSYRLYSRDRGRWGWKDRGOEVIISGNEDEPGVLDLGPVW ARLYRYPADRVMSVCLRVLYGCLWRDGLLSYAPQWYLYSEAVYLNDSYDGH TVGGLQGLGLADGVVGLDDFRKQELRVWPGDYVWNSNHSSFGSYVMEFEFDR LRAFOAMQVHCNNMHTLGAFLPGVVECFRRGPAWEGEPHNLGNLGDPRARV NTFSSLELEPRGQOPVAKAGSPAILIGCLVAIILLILLIALLMLRHLWRLLSKA ERRLVEELTVHLSVPGCTILINRPGPREPPYOEPRGPNPSPACVPNGSAYSG DYMEPEKGAELLPQNSVPHYAADIIVTLOGVTGNTYAVALPFGVGDGPPRV DFRSRLEKLEKIMSLKDNILRLGVCVQDDPLCMITDYMNGDLNQFLSHOLE TKNAENDEKIMSLKDNILRLGVCVQDDPLCMITDYMNGDLNQFLSHOLE DKAAEAGPDGQAAGTPTSYPMLLHVAQAISGMRYLATLNFVRLDRLATNCLVGEN FTIKADFGMSRNLVAGDYRVQGRAVLPIRMWAECLTMCKETASDVWAFGYTLWE VMLCRAQPFGLTDEQVIEAGFEFFRQGVLSRPPACQQLYELMLRCSRESE ORPPFSQLHRLAEDALNTV"

## sig\_peptide

336..423

## gene

337..2967

## CDS

337..2967

/gene="EDDR1"

/standard\_name="epithelial discoidin domain receptor"

/product="receptor tyrosine kinase"

/protein\_id="CAA82335.1"

/db\_xref="PID:g732800"

/db\_xref="GI:732800"

/db\_xref="SWISS-PROT:Q08345"

/translation="MGPEALSLLLLLLVASGDADMKGHFPAKRYALGMDRTIPD SDIASSSSDSTAARHSRLSSDGDGAWCPAGSVFPEEYLOVDLQRLHLVALVGT QGRHAGGLGKEFSYRLYSRDRGRWGWKDRGOEVIISGNEDEPGVLDLGPVW ARLYRYPADRVMSVCLRVLYGCLWRDGLLSYAPQWYLYSEAVYLNDSYDGH TVGGLQGLGLADGVVGLDDFRKQELRVWPGDYVWNSNHSSFGSYVMEFEFDR LRAFOAMQVHCNNMHTLGAFLPGVVECFRRGPAWEGEPHNLGNLGDPRARV NTFSSLELEPRGQOPVAKAGSPAILIGCLVAIILLILLIALLMLRHLWRLLSKA ERRLVEELTVHLSVPGCTILINRPGPREPPYOEPRGPNPSPACVPNGSAYSG DYMEPEKGAELLPQNSVPHYAADIIVTLOGVTGNTYAVALPFGVGDGPPRV DFRSRLEKLEKIMSLKDNILRLGVCVQDDPLCMITDYMNGDLNQFLSHOLE TKNAENDEKIMSLKDNILRLGVCVQDDPLCMITDYMNGDLNQFLSHOLE DKAAEAGPDGQAAGTPTSYPMLLHVAQAISGMRYLATLNFVRLDRLATNCLVGEN FTIKADFGMSRNLVAGDYRVQGRAVLPIRMWAECLTMCKETASDVWAFGYTLWE VMLCRAQPFGLTDEQVIEAGFEFFRQGVLSRPPACQQLYELMLRCSRESE ORPPFSQLHRLAEDALNTV"

polyA\_signal

3807

polyA\_site

3828

BASE COUNT

750 a 1123 c 1126 g 842 t

ORIGIN

1

GATGCTGACATGAAGGACATTTTGTATCTGCTGCAAGTGCCTATGCTGCGGATGCGAG 60

Db 391 GATGCTGACATGAAGGACATTTTGTATCTGCTGCAAGTGCCTATGCTGCGGATGCGAG 450

Query Match 99.9%; Score 1195.4; DB 10; Length 3841;

Best Local Similarity 99.9%; Pred. No. 1.5e-248;

Matches 1196; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GATGCTGACATGAAGGACATTTTGTATCTGCTGCAAGTGCCTATGCTGCGGATGCGAG 60

Db 391 GATGCTGACATGAAGGACATTTTGTATCTGCTGCAAGTGCCTATGCTGCGGATGCGAG 450

QY 61 GACCGGACCATCCAGACAGTACATCTCTGCTTCAGCTCCTGTCAGATTCACATGCC 120  
 DB 451 GACCGGACCATCCAGACAGTACATCTCTGCTTCAGCTCCTGTCAGATTCACATGCC 510  
 QY 121 GCGCGCACAGAGTTGGAGAGCAGTACGCGGATGGGCTGGTGGCCCGCAGGTCG 180  
 DB 511 GCGCGCACAGAGTTGGAGAGCAGTACGCGGATGGGCTGGTGGCCCGCAGGTCG 570  
 QY 181 GTGTTTCCCAAGAGAGAGTACTTGCAGGTGGATCTACACGACTGCACCTGGTGCT 240  
 DB 571 GTGTTTCCCAAGAGAGAGTACTTGCAGGTGGATCTACACGACTGCACCTGGTGCT 630  
 QY 241 CTGGTGGGACACGAGGACGATCGCGGGCCCTGGGCAAGAGATTTCTCCGAGGTAC 300  
 DB 631 CTGGTGGGACACGAGGACGATCGCGGGCCCTGGGCAAGAGATTTCTCCGAGGTAC 690  
 QY 301 CGGCTGCGTTACTCCCGGATGGTCCGCTGGATGGCTGGAAGCCGCTGGGTCAG 360  
 DB 691 CGGCTGCGTTACTCCCGGATGGTCCGCTGGATGGCTGGAAGCCGCTGGGTCAG 750  
 QY 361 GAGGTGATCTCAGCAATGAGGACCTCTGAGGAGTGGTCTGAAGGACCTTGGGCCGCC 420  
 DB 751 GAGGTGATCTCAGCAATGAGGACCTCTGAGGAGTGGTCTGAAGGACCTTGGGCCGCC 810  
 QY 421 ATGTTGCCCGACTGGTCTGCTTACCCCGGGCTGACCGGGTCTAGGCGTCTGTCTG 480  
 DB 811 ATGTTGCCCGACTGGTCTGCTTACCCCGGGCTGACCGGGTCTAGGCGTCTGTCTG 870  
 QY 481 CGGTTAGAGCTCTATGGCTGCTCTGGAGGATGAGTCTCTTACACCGCCCTGTG 540  
 DB 871 CGGTTAGAGCTCTATGGCTGCTCTGGAGGATGAGTCTCTTACACCGCCCTGTG 930  
 QY 541 GGGCAGACAATGATTTATCTGAGGCGCTGTACCTCAACGACTCCACCTATGACGGACAT 600  
 DB 931 GGGCAGACAATGATTTATCTGAGGCGCTGTACCTCAACGACTCCACCTATGACGGACAT 990  
 QY 601 ACCGTGGGGGACTCGAGTATGGGGTCTGGGCCAGCTGGCAGATGCTGGTGGGCTG 660  
 DB 991 ACCGTGGGGGACTCGAGTATGGGGTCTGGGCCAGCTGGCAGATGCTGGTGGGCTG 1050  
 QY 661 GATGACTTTAGGAAGAGTCAGAGCTCGGGTCTGGCCAGCTGATGATGATGGGATG 720  
 DB 1051 GATGACTTTAGGAAGAGTCAGAGCTCGGGTCTGGCCAGCTGATGATGATGGGATG 1110  
 QY 721 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGGAGTTGAGTTGACCGGCTGAGG 780  
 DB 1111 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGGAGTTGAGTTGACCGGCTGAGG 1170  
 QY 781 GCTTCCAGGCTATGAGTCCACTGTATACACATGACACGCTGGGAGCCGCTGTGCT 840  
 DB 1171 GCTTCCAGGCTATGAGTCCACTGTATACACATGACACGCTGGGAGCCGCTGTGCT 1230  
 QY 841 GCGGGGTGGATGTCGCTTCGGCGTGGCCCTGCATGGCTGGGAGGGAGCCCATG 900  
 DB 1231 GCGGGGTGGATGTCGCTTCGGCGTGGCCCTGCATGGCTGGGAGGGAGCCCATG 1290  
 QY 901 CGCCACAACCTAGGGGGAACCTGGGGACCCAGAGCCGCGGCTGCTCAGTGCCCTT 960  
 DB 1291 CGCCACAACCTAGGGGGAACCTGGGGACCCAGAGCCGCGGCTGCTCAGTGCCCTT 1350  
 QY 961 GCGGGCGTGGTGGCTTCTGTCAGTGGCGCTTCTCTTGGGGGCCCTGGTTACTC 1020  
 DB 1351 GCGGGCGTGGTGGCTTCTGTCAGTGGCGCTTCTCTTGGGGGCCCTGGTTACTC 1410  
 QY 1021 TTCACGGAATCTCTTCTATCTCTGATGTGTGAACAATTCCTCCGGGACTGGAGGC 1080  
 DB 1411 TTCACGGAATCTCTTCTATCTCTGATGTGTGAACAATTCCTCCGGGACTGGAGGC 1470  
 QY 1081 ACCTTCCCGGACGCCCTCTGGTGGCGGCTGGCCACCTCCACCACTTCAGCAGCTTG 1140  
 DB 1471 ACCTTCCCGGACGCCCTCTGGTGGCGGCTGGCCACCTCCACCACTTCAGCAGCTTG 1530

QY 1141 GAGCTGGAGCCAGAGCCAGCAGCCCGTGGCCAGCCCGGAGGAGCCCGACCGCC 1197  
 DB 1531 GAGCTGGAGCCAGAGCCAGCAGCCCGTGGCCAGCCCGGAGGAGCCCGACCGCC 1587

RESULT 5  
 HUMCAKA  
 LOCUS  
 DEFINITION Homo sapiens Cak receptor kinase mRNA, complete cds.  
 ACCESSION U57508  
 NID 91160924  
 VERSION 157508.1 GI:1160924  
 KEYWORDS receptor kinase.  
 SOURCE Homo sapiens lung cDNA to mRNA.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 3692)  
 AUTHORS Perez, J.L., Jing, S.Q. and Wong, T.W.  
 TITLE Identification of two isoforms of the Cak receptor kinase that are coexpressed in breast tumor cell lines  
 JOURNAL Oncogene 12 (7), 1469-1477 (1996)  
 MEDLINE 96204002  
 FEATURES  
 Location/Qualifiers  
 1..3692  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /tissue\_type="lung"  
 /note="(vector lambda gt10)"  
 193..2823  
 /gene="Cak"  
 193..2823  
 /gene="Cak"  
 /codon\_start=1  
 /product="receptor kinase"  
 /protein\_id="AAB05208.1"  
 /db\_xref="PID:g1160925"  
 /db\_xref="GI:1160925"  
 /translation="MGPEALSSLLLLVASGDADMGKGFDPKCRKRYALGMODRTIPD  
 SDISASSWSDSSTAARHSRLSSDDGAWCPAGSVFPKEEYLVQDLQRLHLVALVGT  
 QRHAGLGKREFRSIRLRYSDRRWRWGWKDRWQVEISGNEDEGVVLLKDLGPPMV  
 AKLRFYPRADRVMSVLRVLYGLWRDGLSTAPVGTMYLSEAYLLNDSTYDGH  
 TVGGLQGLQGLADVGLDDFRKSOELRWPMGYDYVGSWNSFSYSGYVMEFEFDR  
 LKFAQAMVHCNMHTLGLARLPQGVCEFRFRGPMAMWEGEPMRNLHGKPLDPRARV  
 SVPLGRVARTLQCFELFAGPWLFLSEISFLSDVYVNSSPALGGTFPPAPWPGPPPP  
 TNFSLELETPRQGPVAKAEGSPFALLGLCLVAIIILLLLIALLMLRHLRLLSKA  
 ERRLLEELTVHLSVPGDTILINRPGRPPPPYOEPRGNPNPHSPACVPNGSAYSG  
 DYMEPEKPGAPLLPPPPQNSVPHYAADIIVTQGVGTGNTYAVPALPGVAGDPPRV  
 DPPRSLRFKEKLGEGFCEVHLCEDVSPQDLVSLDFPLNVRKGNHPLKLVKILPDA  
 TKNARNDLFKEVKIMSLKDPNIIRLLGCVQDDPLCMITDYMENGLNQLSAHOLE  
 DKAAGAPGDGAOAGPTISYPMLLHVAQIASGNRYLATLNFVHRDLATNRLNGEN  
 FTIKTADGMSRLNYAGDIYRVQGRAVLPIRMMAWECILMGKFTTASDVWAFGVTLWE  
 QMPLCRAPFGQLTDEQVIEAGEEFRDGRQVYLSRPPACPGQLYELMLRCWRESE  
 QRPFSQLRHLAEALNTV"  
 BASE COUNT 704 a 1127 c 1102 g 759 t  
 ORIGIN

Query Match 99.9%; Score 1195.4; DB 10; Length 3692;  
 Best Local Similarity 99.9%; Pred. No. 1.5e-248;  
 Matches 1196; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GATGCTGACATGAAGGACATTTTGATCTGCTGCCAAGTGGCGCTATGCCCTGGGCATGCAG 50  
 DB 247 GATGCTGACATGAAGGACATTTTGATCTGCTGCCAAGTGGCGCTATGCCCTGGGCATGCAG 306  
 QY 61 GACCGGACCATCCAGACAGTACATCTCTGCTTCAGTCTCTGCTGGTCAGATTCCTGCTGCC 120  
 DB 307 GACCGGACCATCCAGACAGTACATCTCTGCTTCAGTCTCTGCTGGTCAGATTCCTGCTGCC 366  
 QY 121 GCGCGCCACAGCAGGTGGAGAGCAGTACGCGGGATGGGCGCTGGTGGCCCGCAGCGGTG 180  
 DB 367 GCGCGCCACAGCAGGTGGAGAGCAGTACGCGGGATGGGCGCTGGTGGCCCGCAGCGGTG 426

QY 181 GTGTTTCCCAAGGAGGAGTACTTGCAGGTGATCTACAACGACTGCACCTGGTGGCT 240  
DB 427 GTGTTTCCCAAGGAGGAGTACTTGCAGGTGATCTACAACGACTGCACCTGGTGGCT 486  
QY 241 CTGTTGGGACCCAGGACGCGCATGCCGGGGCTGGCAAGGAGTTCTCCCGAGCTAC 300  
DB 487 CTGTTGGGACCCAGGACGCGCATGCCGGGGCTGGCAAGGAGTTCTCCCGAGCTAC 546  
QY 301 CGGTGCTTACTCCCGGATGTCGCGCTGGATGGCTGGAAGACCGCTGGGTGAG 360  
DB 547 CGGTGCTTACTCCCGGATGTCGCGCTGGATGGCTGGAAGACCGCTGGGTGAG 606  
QY 361 GAGGTGATCTCAGGCAATGAGGACCCCTGAGGAGTGTGTGAAGGACCTTGGGCCCCC 420  
DB 607 GAGGTGATCTCAGGCAATGAGGACCCCTGAGGAGTGTGTGAAGGACCTTGGGCCCCC 666  
QY 421 ATGTTGCCCGACTGTTGCTTCTACCCCGGCTGACCGGTCATGAGGCTGCTGCTG 480  
DB 667 ATGTTGCCCGACTGTTGCTTCTACCCCGGCTGACCGGTCATGAGGCTGCTGCTG 726  
QY 481 CGGTGAGGCTCTATGCTGCTCTGGAGGATGAGTCTGCTTACACCGCCCTGTG 540  
DB 727 CGGTGAGGCTCTATGCTGCTCTGGAGGATGAGTCTGCTTACACCGCCCTGTG 786  
QY 541 GGGCAGACAATGATTAATCTAGGCGCTGATCTCAACGACTCCACCTATGACGACAT 600  
DB 787 GGGCAGACAATGATTAATCTAGGCGCTGATCTCAACGACTCCACCTATGACGACAT 846  
QY 601 ACCGTGGCGGACCTCCAGTATGGGGCTGCGGCGAGCTGGCAGATGTGTGGGGCTG 660  
DB 847 ACCGTGGCGGACCTCCAGTATGGGGCTGCGGCGAGCTGGCAGATGTGTGGGGCTG 906  
QY 661 GATGACTTTAGGAAGAGTACGAGCTGCGGCTGCGGCGAGCTATGACTATGTTGGATGG 720  
DB 907 GATGACTTTAGGAAGAGTACGAGCTGCGGCTGCGGCGAGCTATGACTATGTTGGATGG 966  
QY 721 AGCAACACAGCTTCTCAGTGGCTATGTGGAGATGAGTTGAGTTGACCGGCTGAGG 780  
DB 967 AGCAACACAGCTTCTCAGTGGCTATGTGGAGATGAGTTGAGTTGACCGGCTGAGG 1026  
QY 781 GCTTCCAGGCTATGCAAGTCCACTGTAAACAACATGACAGCTGGGAGCCGCTGCTGCT 840  
DB 1027 GCTTCCAGGCTATGCAAGTCCACTGTAAACAACATGACAGCTGGGAGCCGCTGCTGCT 1086  
QY 841 GCGGGGTGGAATGTGCTTCCGGCTGCGGCTGCCATGGCTGGGAGGGAGGCCCATG 900  
DB 1087 GCGGGGTGGAATGTGCTTCCGGCTGCGGCTGCCATGGCTGGGAGGGAGGCCCATG 1146  
QY 901 GCGCAACCTAGGGGCAACCTGGGGACCCAGAGCCGGGCTGTCTAGTGGCCCTT 960  
DB 1147 GCGCAACCTAGGGGCAACCTGGGGACCCAGAGCCGGGCTGTCTAGTGGCCCTT 1206  
QY 961 GCGCGCGTGGCTGCTTCTGAGTGGCGCTTCTTTCGCGGCGCTGCTGCTGCTGCTGCT 1020  
DB 1207 GCGCGCGTGGCTGCTTCTGAGTGGCGCTTCTTTCGCGGCGCTGCTGCTGCTGCTGCT 1266  
QY 1021 TTCAGCAAAATCTCCTTCTGATGTGTGAACAATTCCTTCCGGGACTGGGAGGC 1080  
DB 1267 TTCAGCAAAATCTCCTTCTGATGTGTGAACAATTCCTTCCGGGACTGGGAGGC 1326  
QY 1081 ACCTTCCCGCAGCCCTTGTGGCGCTTGGCCACCTCCACCACTTCCAGCAGCTT 1140  
DB 1327 ACCTTCCCGCAGCCCTTGTGGCGCTTGGCCACCTCCACCACTTCCAGCAGCTT 1386  
QY 1141 GAGCTGGAGCCAGAGCCCTGCTGGCCAGCCCGAGGGAGCCGACCCGCC 1197  
DB 1387 GAGCTGGAGCCAGAGCCCTGCTGGCCAGCCCGAGGGAGCCGACCCGCC 1443  
RESULT 6  
I68738 3962 bp DNA PAT 23-DEC-1997  
LOCUS I68738

DEFINITION Sequence 1 from patent US 5677144.  
ACCESSION I68738  
NID 92830860  
VERSION I68738.1 GI:2830860  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 3962)  
AUTHORS Ullrich,A. and Alves,F.HildegardElisabeth.  
TITLE Recombinant DNA encoding CK 2, a receptor tyrosine kinase  
JOURNAL Patent: US 5677144-A 1 14-OCT-1997;  
FEATURES Location/Qualifiers  
source 1..3962  
BASE COUNT 735 a 1234 c 1182 g 811 t  
ORIGIN  
Query Match 99.6%; Score 1192.2; DB 5; Length 3962;  
Best Local Similarity 99.7%; Pred. No. 7.2e-248;  
Matches 1194; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 GATGCTGACATGAAGGACATTTTGCATCTGCCAAGTCCCGCTATGCCCTGGGCATGCAG 60  
DB 375 GATGCTGACATGAAGGACATTTTGCATCTGCCAAGTCCCGCTATGCCCTGGGCATGCAG 434  
QY 61 GACCGGACCATCCAGACAGTGCATCTCTGCTTCCAGCTCTCTGCTCAGATTCCACTGCC 120  
DB 435 GACCGGACCATCCAGACAGTGCATCTCTGCTTCCAGCTCTCTGCTCAGATTCCACTGCC 494  
QY 121 GCGCGCCACAGCAGGTTGGAGAGCAGTACGCGGGATGGGGCTTGGTGCCTGGCGAGGTCG 180  
DB 495 GCGCGCCACAGCAGGTTGGAGAGCAGTACGCGGGATGGGGCTTGGTGCCTGGCGAGGTCG 554  
QY 181 GTGTTTCCCAAGGAGGAGTACTTGCAGTGGATCTACAACGACTGCACCTGGTGGCT 240  
DB 555 GTGTTTCCCAAGGAGGAGTACTTGCAGTGGATCTACAACGACTGCACCTGGTGGCT 614  
QY 241 CTGTTGGGACCCAGGAGCGCATGCCGGGGCTTGGCAAGGAGTTCTCCCGAGCTAC 300  
DB 615 CTGTTGGGACCCAGGAGCGCATGCCGGGGCTTGGCAAGGAGTTCTCCCGAGCTAC 674  
QY 301 CGGTGCTTACTCCCGGATGGTCCGCTGGATGGCTGGAAGGACCGCTGGGGTCAG 360  
DB 675 CGGTGCTTACTCCCGGATGGTCCGCTGGATGGCTGGAAGGACCGCTGGGGTCAG 734  
QY 361 GAGGTGATCTCAGGCAATGAGGACCCCTGAGGAGTGTGTGAAGGACCTTGGGCCCCC 420  
DB 735 GAGGTGATCTCAGGCAATGAGGACCCCTGAGGAGTGTGTGAAGGACCTTGGGCCCCC 794  
QY 421 ATGTTTCCCGACTGGTTCGCTTCTACCCCGGCTGACCGGCTCATGAGGCTCTGTCTG 480  
DB 795 ATGTTTCCCGACTGGTTCGCTTCTACCCCGGCTGACCGGCTCATGAGTCTGTCTG 854  
QY 481 CGGTGAGACTCTATGGCTGCTCTGGAGGATGAGCTCTGCTTACACCGCCCTGTG 540  
DB 855 CGGTGAGACTCTATGGCTGCTCTGGAGGATGAGCTCTGCTTACACCGCCCTGTG 914  
QY 541 GGGCAGACAATGATTAATCTGAGGCGCTGATCTCAACGACTCCACCTATGACGACAT 600  
DB 915 GGGCAGACAATGATTAATCTGAGGCGCTGATCTCAACGACTCCACCTATGACGACAT 974  
QY 601 ACCGTGGCGGACTGCAGTATGGGGCTTGGGCCAGCTGGCAGATGGTGTGGGGCTG 660  
DB 975 ACCGTGGCGGACTGCAGTATGGGGCTTGGGCCAGCTGGCAGATGGTGTGGGGCTG 1034  
QY 661 GATGACTTTAGGAAGAGTGCAGGAGTGGGGCTTGGCCAGGCTATGACTATGTGGATGG 720  
DB 1035 GATGACTTTAGGAAGAGTGCAGGAGTGGGGCTTGGCCAGGCTATGACTATGTGGATGG 1094  
QY 721 AGCAACACAGCTTCTCAGTGGCTATGAGAGTGGAGTTTGTGTTGACCGGCTGAGG 780

```

Db 1095 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGGAGTTTGAGTTTGACCGGCTGAGG 1154
Qy 781 GCCTTCAGGCTATGAGGTCACCTGTAACAAATGACACAGCTGGAGCCGCTGCTGCT 840
Db 1155 GCCTTCAGGCTATGAGGTCACCTGTAACAAATGACACAGCTGGAGCCGCTGCTGCT 1214
Qy 841 GCGGGGTGAATGTCCTTCGCGGTGGCCCTGCGCATGCGCTGGAGGGGAGCCCATG 900
Db 1215 GCGGGGTGAATGTCCTTCGCGGTGGCCCTGCGCATGCGCTGGAGGGGAGCCCATG 1274
Qy 901 CGCCACAACCTAGGGGCAACCTGGGGAGCCCGAGAGCCGCGCTCTCAGTGGCCCTT 960
Db 1275 CGCCACAACCTAGGGGCAACCTGGGGAGCCCGAGAGCCGCGCTCTCAGTGGCCCTT 1334
Qy 961 GCGGGCGTGGCTGCTTCCTTCGAGTGGCGCTTCTCTTTTCGGGGCGCTGTTACTC 1020
Db 1335 GCGGGCGTGGCTGCTTCCTTCGAGTGGCGCTTCTCTTTTCGGGGCGCTGTTACTC 1394
Qy 1021 TTCAGCAAAATCTCTTCATCTCTGATGGTGAACAAATTCCTTCGCGCACTGGAGGC 1080
Db 1395 TTCAGCAAAATCTCTTCATCTCTGATGGTGAACAAATTCCTTCGCGCACTGGAGGC 1454
Qy 1081 ACCTTCGCGCAGCCCTCGTGGTGGCGCTGGCCCACTCCCAACCAATTCCTTCGCGCAGCTTG 1140
Db 1455 ACCTTCGCGCAGCCCTCGTGGTGGCGCTGGCCCACTCCCAACCAATTCCTTCGCGCAGCTTG 1514
Qy 1141 GAGCTGAGCCAGAGGCCAGCCGCTGGCCCAAGCCGAGGGAGCCCGCCG 1197
Db 1515 GAGCTGAGCCAGAGGCCAGCCGCTGGCCCAAGCCGAGGGAGCCCGCCG 1571

RESULT 7
HSTRKE HSTRKE 3554 bp mRNA PRI 23-JAN-1997
LOCUS H. sapiens TRK E mRNA.
DEFINITION X74979
ACCESSION 9400462
VERSION X74979.1 GI:400462
KEYWORDS trk E gene; Trk E protein.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3554)
AUTHORS Di Marco, E., Cutuli, N., Guerra, L., Cancedda, R. and De Luca, M.
TITLE Molecular cloning of trkE, a novel trk-related putative tyrosine
kinase receptor isolated from normal human keratinocytes and widely
expressed by normal human tissues
J. Biol. Chem. 268 (32), 24290-24295 (1993)
JOURNAL 9403265
MEDLINE 2 (bases 1 to 3554)
REFERENCE De Luca, M.
AUTHORS Direct Submission
TITLE Submitted (08-SEP-1993) M. De Luca, Istituto Nazionale per la
JOURNAL Ricerca sul Cancro, Lab Differenzamento Cellulare, Viale
Benedetto XV, 10 - 16132 Genova, ITALY
Location/Qualifiers
1. .3554
/organism="Homo sapiens"
/db_xref="taxon:9606"
/dev_stage="germ line, fetal /adult"
/tissue_type="brain/keratinocytes"
46. .2676
/gene="trke"
/codon_start=1
/product="trke"
/protein_id="CAA52915.1"
/db_xref="PID:9400463"
/db_xref="GI:400463"
/db_xref="SWISS-PROT:008345"
/translation="MGPEALSSLLLLLVASGDADMKHFDPAKCRVALGMQDRTIPD"

```

```

SDIASSSWSDSTAARHSRLSSDDGAMCPAGSVFPKEEYLQVDLQRLHLVALVGT
QGRHAGLGKEFSRSLRYSDRGRWMMGHKDRWQEVISGNEDEPGVYLKDLGPPMV
ARLVFPYPRADRVMSVCLRVLLGCLWRDGLSTAPVGTMYLSAVYVEMFEFDR
TVGGLQYGGGLQGLADGVLDLDFKSOELRWPGYDVGWNSHSFSGVYVEMFEFDR
LRAFQMVHCHNNHTLGLARLPGLGVEFRFRGPMAMWESSPMRHNGLGNDPPRARV
SVPLGQVAREFLQCFRLFAGPWLIFSEISFISDVWNSSPALGCTTPAPWPPGPP
TNFSSLELEPRGQOPVAKAGSPTAILIGCLVALILLLLI IALMLRHLWRLLSKA
ERVLEBELTVHLSVPGDITLINRPGPREPPYQPRPGNPPHAPCPVNGSAYSG
DYMEPEKPAIPLPPPPQNSVPHYAEADIVTLOGVTGGNTYAPALPPGAGGPPRV
DPSRDLRFKEKLGEGFGEVHUCEVDSDPLVSLDFPLNRKGHLLVAVKTLRDPDA
TKNARDLFEKIMSKRLKDPNIRLLGVCOVDDPLCMITDYMENGLNVLQSAHLE
DKAAGAPGDGAQOQFTISYPMLLHVAQIASGMRYLATLNFVHRLATRNCLGLEN
FTIKADFGSMRLNYAGDYRVQGRVLPTRWMAWECILMKFTTASDVYAFQVTLWE
VLMCRAOPFGQLTDQVIEENAGEFFRDQGRQVYLSRPPACPOGLYELMLRCSRESE
QRPPEQLHREFAEDALNTV"
BASE COUNT 682 a 1064 c 1065 g 743 t
ORIGIN
Query Match 99.6%; Score 1192.2; DB 10; Length 3554;
Best Local Similarity 99.7%; Pred. No. 7.3e-248;
Matches 1194; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 GATGCTGACATGAAGGACATTTTGTATCTGCTGCAAGTGGCGCTATGCTGGGCGATCGAG 60
Db 100 GATGCTGACATGAAGGACATTTTGTATCTGCTGCAAGTGGCGCTATGCTGGGCGATCGAG 159
Qy 61 GACCGGACCATCCAGACAGTACATCTGCTTCCAGCTCCTGGTTCAGATTCACATGCC 120
Db 160 GACCGGACCATCCAGACAGTACATCTGCTTCCAGCTCCTGGTTCAGATTCACATGCC 219
Qy 121 GCCCGCCACAGCAGTGGAGAGCAGTGACGGGATGGGCGCTGGCGCCCGCAGGCTCG 180
Db 220 GCCCGCCACAGCAGTGGAGAGCAGTGACGGGATGGGCGCTGGCGCCCGCAGGCTCG 279
Qy 181 GTGTTTCCCAAGGAGGAGGAGTACTTTCAGTGGATCTACAACGACTSCACCTTGGTGGCT 240
Db 280 GTGTTTCCCAAGGAGGAGGAGTACTTTCAGTGGATCTACAACGACTSCACCTTGGTGGCT 339
Qy 241 CTGCTGGGCCACCCAGGACGCGCATGCCGGGGCTGGGCAAGAGTTCCTCCCGAGGATAC 300
Db 340 CTGCTGGGCCACCCAGGACGCGCATGCCGGGGCTGGGCAAGAGTTCCTCCCGAGGATAC 399
Qy 301 CGGTGGCTTACTCCCGGGATGGTGGCGCTGGATGGCTGGAAGGACCCCTGGGGTCAG 360
Db 400 CGGTGGCTTACTCCCGGGATGGTGGCGCTGGATGGCTGGAAGGACCCCTGGGGTCAG 459
Qy 361 GAGTGATCTCAGCAATAGGACCCCTGAGGGAGTGGTGTCTGAAGGACCTTGGGGCCCCC 420
Db 460 GAGTGATCTCAGCAATAGGACCCCTGAGGGAGTGGTGTCTGAAGGACCTTGGGGCCCCC 519
Qy 421 ATGTTTCCCGACGTTCTGCTTCTACCCCGGGCTGACCGGGCTCATGAGCGTGTCTG 480
Db 520 ATGTTTCCCGACGTTCTGCTTCTACCCCGGGCTGACCGGGCTCATGAGCGTGTCTG 579
Qy 481 CGGTAGAGCTCTATGCTGCTCTGGAGGATGGAGTCTCTTACACCGCCCTGTG 540
Db 580 CGGTAGAGCTCTATGCTGCTCTGGAGGATGGAGTCTCTTACACCGCCCTGTG 639
Qy 541 GGGCAGACATGATTTATCTGAGGCGCTGTACTCTAACGACTCCACCTATGACGGACAT 600
Db 640 GGGCAGACATGATTTATCTGAGGCGCTGTACTCTAACGACTCCACCTATGACGGACAT 699
Qy 601 ACCGTGGCGGACGTGACGATGGGGTCTGGGCCAGCTGGCAGATGGTGTGGGGCTG 660
Db 700 ACCGTGGCGGACGTGACGATGGGGTCTGGGCCAGCTGGCAGATGGTGTGGGGCTG 759
Qy 661 GATGACTTTAGGAAGAGTACAGGAGCTCGGGCTCTGGCCAGGCTATGACTATGTGGATGG 720
Db 760 GATGACTTTAGGAAGAGTACAGGAGCTCGGGCTCTGGCCAGGCTATGACTATGTGGATGG 819
Qy 721 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGAGTTTGAGTTTGACCGGCTGAGG 780

```

```

Db 820 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGGAGTTGAGTTGACCCGCTGAGG 879
Qy 781 GCCTCCAGGCTATCAGGTCCACCTGTAAACATGACACAGCTGGAGCCGCTGCTGCT 840
Db 880 GCCTCCAGGCTATCAGGTCCACCTGTAAACATGACACAGCTGGAGCCGCTGCTGCT 939
Qy 841 GCGGGGTGGAATGTCGTTCCGGCGTGGCCCTGTCATGTCGTCGTCGTCGTCGTCGTC 900
Db 940 GCGGGGTGGAATGTCGTTCCGGCGTGGCCCTGTCATGTCGTCGTCGTCGTCGTCGTC 999
Qy 901 GCGCACACCTAGGGGCAACCTCGGGGACCCAGAGCCGCGCTGTCAGTGCCTT 960
Db 1000 GCGCACACCTAGGGGCAACCTCGGGGACCCAGAGCCGCGCTGTCAGTGCCTT 1059
Qy 961 GCGGGCGTGGCTGCTTCTGTCAGTGGCCCTTCTTCTTTCGGGGCCCTGTTACTC 1020
Db 1060 GCGGGCGTGGCTGCTTCTGTCAGTGGCCCTTCTTCTTTCGGGGCCCTGTTACTC 1119
Qy 1021 TTACGGAATCTCTTCTATCTCTGATGTGTGAACAATTCCTTCCGGCACTGGAGGC 1080
Db 1120 TTACGGAATCTCTTCTATCTCTGATGTGTGAACAATTCCTTCCGGCACTGGAGGC 1179
Qy 1081 ACCTTCCGCGAGCCCTTGGTGGCCCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1140
Db 1180 ACCTTCCGCGAGCCCTTGGTGGCCCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 1239
Qy 1141 GAGCTGGAGCCAGAGCCGCTGGGCAAGCCGCGAGGAGCCGCGAGCCGCTG 1197
Db 1240 GAGCTGGAGCCAGAGCCGCTGGGCAAGCCGCGAGGAGCCGCGAGCCGCTG 1296

RESULT 8
HUMRTK HUMRTK 3738 bp mRNA PRI 27-SEP-1993
LOCUS Homo sapiens receptor tyrosine kinase mRNA, complete cds.
DEFINITION L11315
ACCESSION L11315
NID 9403386
VERSION L11315.1 GI:403386
KEYWORDS receptor tyrosine kinase.
SOURCE Homo sapiens (library: lambda-gt11) term placenta cDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 3738)
AUTHORS Johnson,J.D., Edman,J.C. and Rutter,W.J.
TITLE A receptor tyrosine kinase found in breast carcinoma cells has an
extracellular discoidin I-like domain
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 90, 5677-5681 (1993)
MEDLINE 93296201
COMMENT On Sep 27, 1993 this sequence version replaced gi:307397.
FEATURES
Location/Qualifiers
1..3738
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="term placenta"
/tissue_lib="lambda-gt11"
1..141
142..2883
/standard_name="discoidin I-like domain receptor tyrosine
kinase"
/codon_start=1
/product="receptor tyrosine kinase"
/protein_id="AAA02866.1"
/db_xref="PID:9403387"
/db_xref="GI:403387"
/translation="MGPEALSSLLLLVAGSDADMKGHDPKCRVALGMDRTIPD
SDIASSSWSSTAAHRSRLSSDGDGNCAGSVFPKEEYLVQDVLQVHLVYVGT
QGRHAGLGKESRSYRLRYSRDRWGMKDRMGQEVISGNEDPEGVVLKDLGPPW
ARLVRFYPRADRVNCLRVLYGLWRDLGLSTAPQVQWYMLSEAVYINDSYDGH
TVGQLQGLGLADGVVGLDQFRKSQELRVMPGYDYVGVNSHSFSSGYVEMEFDR
LRAFOAMVHCNMLTGLARLGVGVECFRRGPAWMEGEPHNLGNLGDPPARAV
SVPLGRVARELQCRFLFAGPWLFLSEISFTSDVNVNNSPALGTFPPAPWPPPPP
TNFSSLELEPRQQQPVAKAGSPTAILIGCLVAITILLLLIALLMLHRLHRRLLSKA

```

```

ERRVLEELTVHLSVPGDTILNRPGRPREPPYQEPFRPNRPHSPAPCVNPGSALLL
SNPARYLLLATYAPRPGPPTPAWAKPTNTQAYSGDYMEPEKPGAPLLPPONS
PHYAADITLQGVGTGNTYAVPALPGAVGDPVDFRFLRKEKLGEGFGEV
HLCIEVSPQDLVSLDFPLNVRKHPLLVAVILRPDATTNARNDFLKEVKMSRLKDP
NIRLLGVCVODDPLMTIDYMGNDLNOVLSAHOLEDKAAAGAPGQQAQSGPTISY
PMLLHVAQIASGRMYLATLNFVHRDLATRNCLVGENFTIKIADFGRSNLYAGDYR
VOGRAVLPIRNMWECILMGAFITTSADSWAFGVVWVLMCLRAQPGQLDQDEVIEN
AGEFFRDQGRQVYLSRPPACPGQLYELMLRCWSESEQRPPFSQLHRFLAEDALNTV
2887..3738
3'UTR
polyA_site 3738
BASE COUNT 698 a 1140 c 1123 g 777 t
ORIGIN
Query Match 99.5%; Score 1190.6; DB 9; Length 3738;
Best Local Similarity 99.7%; Pred. No. 1.6e-247;
Matches 1193; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 1 GATGCTGACATGAAGGACATTTTGTATCTGCCAAGTGCCTATGCCCTGGGCATGCAG 50
Db 196 GATGCTGACATGAAGGACATTTTGTATCTGCCAAGTGCCTATGCCCTGGGCATGCAG 255
Qy 61 GACCGACCATCCACACAGTACATCTCTGCTCCAGCTCCTGGTGCAGATTCCTACTGCC 120
Db 256 GACCGACCATCCACACAGTACATCTCTGCTCCAGCTCCTGGTGCAGATTCCTACTGCC 315
Qy 121 GCGCGCCACACAGAGTGGAGAGCAGTACGAGGAGTGGGGGCTGTGGTCCCGCAGGGTGC 180
Db 316 GCGCGCCACACAGAGTGGAGAGCAGTACGAGGAGTGGGGGCTGTGGTCCCGCAGGGTGC 375
Qy 181 GTGTTTCCCAAGGAGGAGTACTTGTGAGTGGATCTACACGACTGCACCTGGTGGCT 240
Db 376 GTGTTTCCCAAGGAGGAGTACTTGTGAGTGGATCTACACGACTGCACCTGGTGGCT 435
Qy 241 CTGTTGGGACCCACAGGAGCGCATGCGGGGCTCGGCAAGGAGTTCCTCCGGAGCTAC 300
Db 436 CTGTTGGGACCCACAGGAGCGCATGCGGGGCTCGGCAAGGAGTTCCTCCGGAGCTAC 495
Qy 301 CGGCTGCTTACTCCCGGGATGTCGCCCTGATGGGTGGAGGACCGCTGGGGTTCAG 360
Db 496 CGGCTGCTTACTCCCGGGATGTCGCCCTGATGGGTGGAGGACCGCTGGGGTTCAG 555
Qy 361 GAGTGTATCTCAGCAATGAGGACCTGAGGAGTGGTCTGAAGGACCTTGGGCCCCC 420
Db 556 GAGTGTATCTCAGCAATGAGGACCTGAGGAGTGGTCTGAAGGACCTTGGGCCCCC 615
Qy 421 ATGTTTCCCGACTGTTCTGCTTCTACCCCGGCTGACCGGCTCATGAGCGTCTGCTG 480
Db 616 ATGTTTCCCGACTGTTCTGCTTCTACCCCGGCTGACCGGCTCATGAGTCTGCTG 675
Qy 481 CGGCTAGAGCTCTATGCTGCTGCTGAGGAGTGGACTCTGCTTACACCGCCCTGTG 540
Db 676 CGGCTAGAGCTCTATGCTGCTGCTGAGGAGTGGACTCTGCTTACACCGCCCTGTG 735
Qy 541 GGGCAGACATGTTATTTCTGAGGCGCTGTTACTCAACGACTCCACCTATGACGACAT 600
Db 736 GGGCAGACATGTTATTTCTGAGGCGCTGTTACTCAACGACTCCACCTATGACGACAT 795
Qy 601 ACCGTGGCGGACTGTCAGTATGGGGTCTGGCCAGCTGGCAGATGGTGTGGGGCTG 660
Db 796 ACCGTGGCGGACTGTCAGTATGGGGTCTGGCCAGCTGGCAGATGGTGTGGGGCTG 855
Qy 661 GATGACATTTAGGAAGAGTACGAGGCTGCGGGTCTGGCCAGGCTATGACTATGTTGGATGG 720
Db 856 GATGACATTTAGGAAGAGTACGAGGCTGCGGGTCTGGCCAGGCTATGACTATGTTGGATGG 915
Qy 721 AGCAACACAGCTTCTCCAGTGGCTATGTGAGATGGATTTGAGTTTGACCGGCTGAGG 780
Db 916 AGCAACACAGCTTCTCCAGTGGCTATGTGAGATGGATTTGAGTTTGACCGGCTGAGG 975
Qy 781 GCGTTCACAGGCTATGACAGGCTCCACTGTAAACATGACACGCTGGGAGCCGCTGCTGCCT 840
Db 976 GCGTTCACAGGCTATGACAGGCTCCACTGTAAACATGACACGCTGGGAGCCGCTGCTGCCT 1035

```

```

QY 841 GCGGGGTGAATGTCCTTCGGCGCTGGCCCTGCGCATGCGCTGGAGGGGAGGCCATG 900
Db 1036 GCGGGGTGAATGTCCTTCGGCGCTGGCCCTGCGCATGCGCTGGAGGGGAGGCCATG 1095
QY 901 CGCCACAACCTAGGGGGCAACCTGGGGAGCCCGAGAGCCCGGGCTGTCTCAGTGCCTTT 960
Db 1096 CGCCACAACCTAGGGGGCAACCTGGGGAGCCCGAGAGCCCGGGCTGTCTCAGTGCCTTT 1155
QY 961 GCGGGCGTGTGCTCGCTTCCTTCGAGTGCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 1020
Db 1156 GCGGGCGTGTGCTCGCTTCCTTCGAGTGCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 1215
QY 1021 TTCAAGCAATCTCTTCATCTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 1080
Db 1216 TTCAAGCAATCTCTTCATCTCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 1275
QY 1081 ACCTTCCTCCGAGCCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 1140
Db 1276 ACCTTCCTCCGAGCCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 1335
QY 1141 GAGCTGAGCCAGAGCCGAGCCGAGCCGAGCCGAGCCGAGCCGAGCCGAGCCGAGCCGAGCC 1197
Db 1336 GAGCTGAGCCAGAGCCGAGCCGAGCCGAGCCGAGCCGAGCCGAGCCGAGCCGAGCCGAGCC 1392

RESULT 9
LOCUS A42378 3754 bp DNA PAT 06-MAR-1997
DEFINITION Sequence 1 from Patent WO9502187.
ACCESSION A42378
NID 92297854
VERSION A42378.1 GI:2297854
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Gusterson,B.A., Crompton,M.R., Mitchell,P.J., Barker,K.T.,
Martindale,J.E., Page,M.J. and Spence,P.
TITLE CELL GROWTH FACTOR RECEPTORS
JOURNAL Patent: WO 9502187-A 1 19-JAN-1995;
CANCER RES INST (GB)
COMMENT Other publication AU 7081094 950206.
FEATURES
source
1..3754
/organism="unidentified"
/db_xref="taxon:32644"
142..2886
/note="unnamed protein product"
/codon_start=1
/protein_id="CAA02688.1"
/db_xref="pid:e305343"
/db_xref="pid:g2297855"
/db_xref="gi:2297855"
/translation="MGPEALSSLLLLLVASGDADMKHFDPAKRYALGMQDRTPD
SDISASSWSSTAAHSRLSSDGDGAWCPAGSVFPKEEYLOVDLQRLHVALVGT
QGRHAGLGKFSRSLRYLRSDGRWMDKQOVISGNEDEPGVVLKDLGPPMV
ARLVFPYPRADRVMSVCLRLVGLWDLGLSVTA PVGOTMYLSEAVLNDSTYDGH
TVGGVGLGLGOLADGVVGLDFKSKQLRWFGYDXVGNNSHSFSSGYVEMEFEDR
LRAQAMQVHCNNHMTLGLRPGVECFRRGPNAMWEGEPMRHNLGNLGDPRARAY
SVPLGVRVARELQCRFLFAGPWSLFEISFDVYNNSSPALGTFPPAPWPPGPP
TNFSSLELPRGPQPVAKAGSPTAILGLVAILILLILLIILALMLRWLRLLSK
AERVLBEELVHLSPQDITLIINRPGPREPPYQPRPNRPNHPSAPCVNPGSALL
LSNPAYRLILLATYARPGRPGPTPAWKPNTQAYSGDYMEPEKFPAGLPPLPPONS
VPHYAEADIVLQGTGNYAVPALPGAVGPPVDFPRSLRKEKLGSGGQGE
VHLEIVDSQDLVSLDFPLNVRKPLILLVAIKLRPADEKRNDFLVEKIMSLRK
PNIIRLGVQVDDPLCMITDYMGNDLNOFLSAHOLEKAAEGAPGQGOAAGQPTIS
YPLLHVAQAQISAGMYRLATLNFVRLDRLATNCLVGNFTIKIADFGMSRLVAGDY
RVQGRVLPPIRMWAEICLMGKFTTASDVWAFVTLWEVLMLCRAOPFGOLTDEQVIE
NAGEFFRDQGRQVYLSRPPACPOGLYELMLRCNSRESEQRPPFSQLHFLAEDALNTV
"
BASE COUNT 712 a 1146 c 1121 g 775 t

```

ORIGIN

Query Match 98.5%; Score 1179.2; DB 5; Length 3754;  
 Best Local Similarity 99.5%; Pred. No. 4.6e-245;  
 Matches 1194; Conservative 0; Mismatches 3; Indels 3; Gaps 1;

```

QY 1 GATGCTCACATGAAGGACATTTTGTGCTGCAAGTGGCGCTATGCCCTGGGCATGCAG 60
Db 196 GATGCTCACATGAAGGACATTTTGTGCTGCAAGTGGCGCTATGCCCTGGGCATGCAG 255
QY 61 GACCGGACCATCCACAGACAGTACATCTCTGCTTCCAGTCTCCCTGGTCCAGATTCAC 120
Db 256 GACCGGACCATCCACAGACAGTACATCTCTGCTTCCAGTCTCCCTGGTCCAGATTCAC 315
QY 121 GCGCGCCACAGCAGGTGGAGACAGTACAGGGGATGGGGCTGGTGGCCCGCAGGGTCG 180
Db 316 GCGCGCCACAGCAGGTGGAGACAGTACAGGGGATGGGGCTGGTGGCCCGCAGGGTCG 375
QY 181 GTGTTTCCCAAGGAGGAGTACTTTCAGGTGATCTACAACGACTGCACCTGGTGGCT 240
Db 376 GTGTTTCCCAAGGAGGAGTACTTTCAGGTGATCTACAACGACTGCACCTGGTGGCT 435
QY 241 CTGCTGGGCAACCCAGGACGGCATGCCGGGGCTGGGCAAGGAGTCTCCCGGAGCTAC 300
Db 436 CTGCTGGGCAACCCAGGACGGCATGCCGGGGCTGGGCAAGGAGTCTCCCGGAGCTAC 495
QY 301 CGGCTGGGTTACTCCCGGGATGTCGCGCGCTGGATGGGTGGAGACCGCTGGGTTCAG 360
Db 496 CGGCTGGGTTACTCCCGGGATGTCGCGCGCTGGATGGGTGGAGACCGCTGGGTTCAG 555
QY 361 GAGGTGATCTCAGGCAATGAGGACCTGAGGGAGTGGTCTGAAGGACCTTGGGGCCCCC 420
Db 556 GAGGTGATCTCAGGCAATGAGGACCTGAGGGAGTGGTCTGAAGGACCTTGGGGCCCCC 615
QY 421 ATGTTGCCCGACTGCTTCCTTACCCCGGGCTGACCGGGTCTAGCGCTCTCTCTG 480
Db 616 ATGTTGCCCGACTGCTTCCTTACCCCGGGCTGACCGGGTCTAGCGCTCTCTCTG 675
QY 481 CGGCTAGAGCTCTATGGTGCCTCTCGAGGGATGGACTCTGTCTTACACCGCCCTGTG 540
Db 676 CGGCTAGAGCTCTATGGTGCCTCTCGAGGGATGGACTCTGTCTTACACCGCCCTGTG 735
QY 541 GGGCAGACAAATGATTTATCTGAGGCGCTGTACTTCAACGACTCCACCTATGACGACAT 600
Db 736 GGGCAGACAAATGATTTATCTGAGGCGCTGTACTTCAACGACTCCACCTATGACGACAT 795
QY 601 ACCGTGGGCGACTGACAGTATGGGGTCTGGGCCAGCTGGCAGATGGTGTGGGGCTG 660
Db 796 ACCGTGGGCGACTGACAGTATGGGGTCTGGGCCAGCTGGCAGATGGTGTGGGGCTG 855
QY 661 GATGACTTTAGGAGAGTACAGGAGCTGGGGTCTGGCCAGGCTATGACTATGTGGATGG 720
Db 856 GATGACTTTAGGAGAGTACAGGAGCTGGGGTCTGGCCAGGCTATGACTATGTGGATGG 915
QY 721 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGAGTTTGGTTTACCGGCTGAGG 780
Db 916 AGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGAGTTTGGTTTACCGGCTGAGG 975
QY 781 GCCTTCCAGGCTATGACAGTTCCTACTGTAACACATGCACACGCTGGAGCCCGCTGCT 840
Db 976 GCCTTCCAGGCTATGACAGTTCCTACTGTAACACATGCACACGCTGGAGCCCGCTGCT 1035
QY 841 GCGGGGTGGAATGTGCTTCGGCGCTGGCCCTGCCATGGCTGGAGGGGAGGCCATG 900
Db 1036 GCGGGGTGGAATGTGCTTCGGCGCTGGCCCTGCCATGGCTGGAGGGGAGGCCATG 1095
QY 901 CGCCACAACCTAGGGGGCAACCTGGGGAGCCCGAGAGCCCGGGCTGTCTCAGTGCCTTT 960
Db 1096 CGCCACAACCTAGGGGGCAACCTGGGGAGCCCGAGAGCCCGGGCTGTCTCAGTGCCTTT 1155
QY 961 GCGGGCGTGTGCTCGCTTCCTTCGAGTGCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 1020

```



```
Db 1156 GCGCGCGTGTGGCTCCTCTTCTGCACTGCGCGCTTCTCTTTCGCGGGGCGCTGTACTC 1215
Qy 1021 TTCAGCAAAATCTCTTCACTCTCTGATGTGGTGAACAAATTCCTCTCCGGCACTGGGAGC 1080
Db 1216 TTCAGCAAAATCTCTTCACTCTCTGATGTGGTGAACAAATTCCTCTCCGGCACTGGGAGC 1275
Qy 1081 ACCTTCCCGCAGCCCGCTGGTGGCGCCCTGGCGCCACTCCCAACCACTTCAGCAGCTTG 1140
Db 1276 ACCTTCCCGCAGCCCGCTGGTGGCGCCCTGGCGCCACTCCCAACCACTTCAGCAGCTTG 1335
Qy 1141 GAGCTGAGCCAGAGCCAGC---AGCCGCTGCCAAGCCCGAGGGAGCCGACCGCC 1197
Db 1336 GAGCTGAGCCAGAGCCAGCAGCCGCGCGTGGCCAAAGCCGAGGGAGCCGACCGCC 1395

RESULT 10
AF026259 3659 bp mRNA ROD 27-OCT-1997
LOCUS Mus musculus receptor-like tyrosine kinase (Nep) mRNA, complete
DEFINITION cds.
ACCESSION AF026259
NID 92564939
VERSION AF026259.1 GI:2564939
KEYWORDS house mouse.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Zerlin,M., Julius,M.A. and Goldfarb,M.
TITLE NEP: a novel receptor-like tyrosine kinase expressed in
proliferating neuroepithelia
Oncogene 8 (10), 2731-2739 (1993)
JOURNAL 93390947
MEDLINE
REFERENCE
AUTHORS Zerlin,M., Julius,M.A. and Goldfarb,M.
TITLE Direct Submission
Submitted (22-SEP-1997) Brookdale Center for Developmental and
Molecular Biology, Mount Sinai School of Medicine, 1 Gustave Levy
Place, New York, NY 10029, USA
JOURNAL
FEATURES
source
1..3659
/organism="Mus musculus"
/strain="129"
/db_xref="taxon:10090"
/cell_type="CCE embryonic stem cells,
differentiation-induced"
1..3659
/gene="Nep"
/101..2836
/gene="Nep"
/notes="NEP; DDRL"
/codon_start=1
/product="receptor-like tyrosine kinase"
/protein_id="A898186.1"
/db_xref="PIR:G2564940"
/db_xref="GI:2564940"
/translation="MGTGTLSSLLLLLLVITGADMKGHFPAKCYALQMQRDRTIP
DSQTSVSSWSSTAAHRSLESDDGAWCPAGVPFKPEEYLVQDLRLRLHLVALVG
TOGRHAGLGCKEFSRSLRYLRSDRRWMDKRWGQEVISGNEPQGVYVLDLPGPM
VARLRYPRADRVMSVCLRVLYGCLWRDGLLSYAPVQTMQLSEVMVHLNDSTYD
GYTAGGQYGLQGLADGVVGLDDFRQSGELRVWPGIDYVGSNSFFPGYVMEFEF
DRUFTVTMGVHNNHMTLARGLPFGVECFKRGPAWEGEPVSRHALGSLGDPAR
AISVPLGGHVRFLQCRFLFAGLPWFSEISFSDVYNDSDTFPPAPWPPGPPPTN
FSSLEPRQGVKPVRAEGSPAILIGLIVAILLLLLLIALMLRHLWRLLSADS
RVLEELTVHLSVPGDPTILNNRGPPEPPYOEPRPGTPPHSAPCVNPQALLSN
PAYRLLLATYARPPGPPPTPWADKPTNTQACSDYMEPEKPGALPPPPNSVPH
YAERDITVLQVGGNTYVAPLPPGAVGDPVPRPRSRFRKLEGGEGQGEVHL
CEVEDPDQLVSSDFPISVHKGHPHLLVAVKILRPDATKNARNDLFKEVIMSLKDPNI
IRLVGCVQDDPLCMITDMENGLNQFLSARLENKATQSGDSTESQGTISYPM
LLHVGAIASMGWRYLATLNEVHRDLATRNCLVGENFTIKIADFGRNRLYAGDYRVQ
GRVPLTRNWAECILMKFTTASDYWAFGVTLWEVLMLCRSOPFQSLDDEQVIEAG
EFFPDQGRQVYLSRPPACPTQLVELMLRCHWSREPEQRPFPALHRLDADLNTV"
```

```
misc_feature 194..658
/genes="Nep"
/notes="encodes discoidin-like domain"
1340..1411
/genes="Nep"
/notes="encodes transmembrane domain"
1922..2809
/genes="Nep"
/notes="encodes tyrosine kinase domain"
BASE COUNT 724 a 1095 c 1075 g 765 t
ORIGIN

Query Match 74.8%; Score 895.6; DB 12; Length 3659;
Best Local Similarity 85.5%; Pred. No. 6.3e-184;
Matches:1026; Conservative 0; Mismatches 159; Indels 15; Gaps 2;

Qy 1 GATGCTGACATGAGGGACATTTTATCTCCCAAGTGCCTATGCTGGGATCG 60
Db 158 GATGCTGACATGAGGGACATTTTATCTCCCAAGTGCCTATGCTGGGATCG 217
Qy 61 GACCGACCATCCAGACAGTACATCTCTCTCCAGCTCCTTGGTCAGATTCCACTGCC 120
Db 218 GACCGACCATCTCTGACAGCATATCTCTGTCCAGCTCCTGTGAGACTCTACCGCT 277
Qy 121 GCCCGCCACAGAGTTGGAGAGAGTACGCGGATGGGCTGGTCCCGCAGGGTCG 180
Db 278 GCCCGCCACAGAGTTGGAGAGAGTACGCGGATGGGCTGGTCCCGCAGGGCT 337
Qy 181 GGTGTTTCCCAAGGAGGAGTACTTGCAGCTGGATCTTACAACGACTGCACCTGTGGCT 240
Db 338 GGTGTTTCCCAAGGAGGAGTACTTGCAGCTGGATCTTACAACGACTGCACCTGTGGCT 397
Qy 241 CTGGTGGGACCCAGGAGCGCATGCCGGGGCCCTGGGCAAGAGTCTTCCCGGAGCTAC 300
Db 398 CTGGTGGGACCCAGGAGCGCATGCCGGGGCCCTGGGCAAGAGTCTTCCCGGAGCTAT 457
Qy 301 CGGCTGCGTTACTCCCGGATGGTCCGCTGGATGGCTGGAAGGACCTGTGGGTCAG 360
Db 458 CGGTTGCGTTACTCCCGAGATGGCGCGCTGGATGGACTGGAAGGACCGCTGGGGACAG 517
Qy 361 GAGGTGATCTCAGGCAATGAGGACCTCAGGAGTGGTGGTGAAGGACTTGGGCCCCC 420
Db 518 GAGGTGATCTCAGGCAATGAGGACCTCAGGAGTGGTGGTGAAGGACTTGGGCCCCC 577
Qy 421 ATGGTTGCCGACTGGTTCGCTTCTTACCCCGGGCTGACCGGGTCATGAGGCTCTGTGTG 480
Db 578 ATGGTTGCCGCTGGTTCGCTTCTTACCCCGGGCTGACCGGGTCATGAGTGTCTGTCTT 637
Qy 481 CGGCTAGAGCTCTATGGCTGCTCTGGAGGATGGACTCTCTTACACCGCCCGCTGTG 540
Db 638 CGGCTAGAGCTCTATGGCTGCTCTGGAGGATGGACTCTCTTACATACAGCCCCCGTG 697
Qy 541 GGGGACAGCAATCTATTTATCTGAGG---CCGTGTACCTCAACGACTCCACCTATGACGGA 597
Db 698 GGGGACAGCAATCTATTTATCTGAGG---CCGTGTACCTCAACGACTCCACCTATGACGGA 757
Qy 598 CATACCTGGGCGGACTGCAGTATGGGGTCTGGGCCAGCTGGCAGAGTGGTGTGGTGGG 657
Db 758 TATACTGTGGAGGGCTGCAGTATGGCGTCTGGGCCAGCTGGCAGAGTGGTGTGGGCG 817
Qy 658 CTGGATGACTTTAGGAAGAGTGCAGGCTGCGGGTCTGGCCAGGCTATGACTATGTGGGA 717
Db 818 CTGGATGATTTTCAGCAAGACGAGGCTGCGGGTCTGGCCAGGCTATGACTATGTGGGA 877
Qy 718 TGGAGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGGAGTTTGGTTTGAACCGGCTG 777
Db 878 TGGAGCAACACAGCTTCTCCAGTGGCTATGTGGAGATGGAGTTTGGTTTGAACCGGCTG 937
Qy 778 AGGGGCTTCCAGGCTATGAGGTCACCTGTACACACAGCTGGGAGCCCGCTGTG 837
Db 938 AGAACCTTCCAGACCATGAGGTCACCTGTACACACAGCTGGGAGCCCGGCTA 997
```









Qy	13	AAGGGACATTTTGATCCTGCCAAGTGCCTGATGCCCTGTGGCATGCAGGACCGGACCATC	72
Db	433	AAAGCTCAGGTTTAATCCAGCTATATCCGCTATCTCTGGCATTCTCAGSAGGCCAGATT	492
Qy	73	CCAGACAGTCACATCTCTGCTCCAGCTCCTCGTCCAGATTCACATGCGCGCCGCCACAGC	132
Db	493	CCAGATGAGGACATCACAGCTTCCAGTCAAGTGGTCAGAGTCCACAGTGCCTCAATAATGGA	552
Qy	133	AGGTTGGAGAGCAGTAGTCAGCGGGATGGGCGCTTGGTCCCCCGCAGGSGTCGGTTTTCCC	192
Db	553	AGGCTGSACTACAAGAAGGGGATGAGCGCTTGGTGCCTTGAGATTCAGAGTGGAACTGAT	612
Qy	193	GA---GGAGGAGTACTTGGCAGGTGGATCTACAACAGCTGCACCTGGTGGCTCTGGTGGCG	249
Db	613	GACCTGAAGGAGTTCTGTCAGATTTGACTTGCACACCCCTCCATTTTATCACCTCTGGTGGG	672
Qy	250	ACCCAGGACGGCATCCGCGGGCGCTGGCGAAGGAGTTCTCCCGGAGCTACCGGCTGGT	309
Db	673	ACCCAGGGCGCCATCGAGAGGTATGCGATCGAGTTTGGCCCCCATGTCAAGATCAAT	732
Qy	310	TACTCCCGGATGGTCCCGCTGGATGGGCTGGAAGACCGCTGGGGTCAGGAGGTGATC	369
Db	733	TACAGTCGGGATGSCACTCGCTGGATCTCTTGGCGGAACCGTCATGGGAACAGGTGCTG	792
Qy	370	TCAGGCAATGAGGACCCCTGAGGAGTGGTGTCTGAAGACCTTTGGGCGCCCCCATGGTGGCC	429
Db	793	GATGGAAATAGTAACCCCTATGACATTTTCCCTAAAGGACTTGGAGCCGCCCATTTGTAGCC	852
Qy	430	CGACTGGTTCGCTTACCCCCGGGCTGACCGGTCATGAGGCTGTGCTCGGSGTAGAG	489
Db	853	AGATTTTCGGGTTCAATTCAGTCCCGACCACTCCATGAATGTGTGATGAGAGTGGAG	912
Qy	490	CTCTATGGCTCCCTCTGGAGGGATGACCTCTGTCTTACACGCCCGCTGTGGGGCAGACA	549
Db	913	CTTTACGCTGTCTCTGGCTAGATGCTTGGTGTCTTACAATGCTCCAGCTGGGCAGAG	972
Qy	550	ATGTAATTTATCTGAGG-----CGTGTACCTCAAGGACTCCAACCTATGACGGACATACC	603
Db	973	TTTGACTCCCTCGAGGTTCCATCATTTATCTCAATGATTTCTGTCTATGATGAGAGCTGTT	1032

Db 2605 AGGCTGACTCAGAAAGGGGATGGAGCCTGGTGCCCTGAGATTCAGTGGAACTGAT 2546  
Qy 193 GA---GGAGGAGTACTTGCAGGTGGATATACAACTGACCTGGTGGCTGCTGGTGGC 249  
Db 2545 GACCTGAAGGAGTTCTGCAGATTGACTTGCACACCTCCATTTATCACTCTGGTGGG 2486  
Qy 250 ACCAGGACCGCATGCGGGGCTGGGCAAGAGTTCTCCGGAGCTACCGGCTGCGT 309  
Db 2485 ACCAGGGGCCCATGAGGAGTTCATGGCATCGAGTTTCCGCCCATGTACAAGATCAAT 2426  
Qy 310 TACTCCCGGATGCTGCGCGCTGGGTGGAGGACCGCTGGGGTACAGAGGTGATC 369  
Db 2425 TACAGTCGGATGACACTCGCTGGATCTTTGGCGGAACCTGATGGGAACAGGTGCTG 2366  
Qy 370 TCAGGCAATGAGCCCTGAGGAGTGGTGTGAAGGACCTTGGGCCCCCATGGTTGCC 429  
Db 2365 GATGAAATAGTAACCCCTATGACATTTCTAAAGGACTTGGAGCCGCCCATTTGTAGCC 2306  
Qy 430 CGACTGGTTCCTTCTACCCCGGGCTGACGGGGTCAAGCGTCTCTGCGGGTAGAG 489  
Db 2305 AGATTGTCCGGTTTCATTCCAGTCAACCGACCACTCCCATGAATGTGTATGAGAGTGGAG 2246  
Qy 490 CTCATGCTGCTGCTGGAGGATGACTCTCTGTCTTACACCGCCCTGTGGGCGACACA 549  
Db 2245 CTTACGGCTGTCTGCTAGATGGCTGGTGTCTTACAATGCTCCAGCTGGGCGACAG 2186  
Qy 550 ATGATTTATCTGAGGC-----CGTGACTCAACGACTCCACTATGACGGACATACC 603  
Db 2185 TTTGACTCCCTGGAGGTTCCATCATTTAATGATGATTTCTGCTATGATGAGCTGTT 2126  
Qy 604 GTGGCGGACTGCAAGTATGGGGGCTGGGCGAGCTGGCAGATGGTGTGGGGCTGGAT 663  
Db 2125 GGATACAGCATGACAGAGGG--CTAGGCCAATTGACCGATGGTGTCTGCGCTGGAC 2069  
Qy 664 GACTTTAGGAGAGTCAAGGACTCGGGCTGCGGAGCTGCGGAGCTATGACTATGCGGATGAGC 723  
Db 2068 GATTTACCCAGCCCATGAATACCACTGTGGCCCGGCTATGACTATGTTGGGCTGGCGG 2009  
Qy 724 AACCAAGCTTCTCCAGTGGCTATGTGGAGATGGAGTTGAGTTGACCGCTGAGGGCC 783  
Db 2008 AACGAGATGCCACCAATGGCTACATGAGATCATGTTGNAATTGACCCCATCAGGAAT 1949  
Qy 784 TTCAGGCTATGCAAGTCCACTGTAACAACATGACACAGCTGGAGGCCCTCTGCCTGGC 843  
Db 1948 TTCATACCATGAAGTCCACTGCAACAACATGTTGCTAAAGGTGTGAAGATCTTTAAG 1889  
Qy 844 GGGGTGGAATGCTGCTCCGCGTGGCCCTGCCATGGCTGGAGGGGGAGCCCATGCC 903  
Db 1888 GAGGTACAGTGTCTACTTCCGCTCTGAA---GCCAGTGTGGGAACCTAATGCCATTTCC 1832  
Qy 904 CACAACCTAGGGGGCAACCTGGGGGACCCAGAGCCGGGCTGCTCAGTGGCCCTTGGC 963  
Db 1831 TTCCCTTGTCTGGATGAGCTCAACCCAGTCTCGGTTGTACGGTGCCTCTCCAC 1772  
Qy 964 GCGGTGTGGCTGCTTCTCAGTGGCGCTTCTCTTTGGGGGCCCTGGTTACTCTTC 1023  
Db 1771 CACCGAATGGCCAGTGGCCATCAAGTGTCAATACCATTTCAGATACCTGGATGATTC 1712  
Qy 1024 AGCGAATCTCCTTCATCTCTGATGTTGGTGAACAATTCCTCCGGGACTGGGAGGCACC 1083  
Db 1711 AGTGAGATCACCTTCCATCATGATGCTGCAATGTACAACAACACTCTGAAGCCCTGCCACC 1652  
Qy 1084 TTCCCGCCAGCCGCC 1098  
Db 1651 TCTCTATGGCACCC 1637

Search completed: November 4, 1999, 07:31:51  
Job time: 6510 sec

